SIGNAL PROCESSING APPARATUS
Background of the Invention
Reference to Prior Related Application

This is a continuation-in-part application of U.S. Patent Application No. 08/132,812 filed October 6, 1993, and entitled "Signal Processing Apparatus" and a continuation-in-part application of U.S. Patent Application No. 08/249,690 filed May 26, 1994 entitled "Signal Processing Apparatus and Method" which is a continuation of U.S. Patent Application No. 07/666,060 filed March 7, 1991, now abandoned.

Field of the Invention

The present invention relates to the field of signal processing. More specifically, the present invention relates to the processing of measured signals, containing a primary signal portion and a secondary signal portion, for the removal or derivation of either the primary or secondary signal portion when little is known about either of these components. More particularly, the present invention relates to modeling the measured signals in a novel way which facilitates minimizing the correlation between the primary signal portion and the secondary signal portion in order to produce a primary and/or secondary signal. The present invention is especially useful for physiological monitoring systems including blood oxygen saturation systems.

Description of the Related Art

Signal processors are typically employed to remove or derive either the primary or secondary signal portion from a composite measured signal including a primary signal portion and a secondary signal portion. For example, a composite signal may contain noise and desirable portions. If the secondary signal portion occupies a different frequency spectrum than the primary signal portion, then conventional filtering techniques such as low pass, band pass, and high pass filtering are available to remove or derive either the primary or the secondary signal portion from the total

signal. Fixed single or multiple notch filters could also be employed if the primary and/or secondary signal portion(s) exist at a fixed frequency(s).

It is often the case that an overlap in frequency spectrum between the primary and secondary signal portions exists. Complicating matters further, the statistical properties of one or both of the primary and secondary signal portions change with time. In such cases, conventional filtering techniques are ineffective in extracting either the primary or secondary signal. If, however, a description of either the primary or secondary signal portion can be derived, correlation canceling, such as adaptive noise canceling, can be employed to remove either the primary or secondary signal portion of the signal isolating the other portion. In other words, given sufficient information about one of the signal portions, that signal portion can be extracted.

Conventional correlation cancelers, such as adaptive noise cancelers, dynamically change their transfer function to adapt to and remove portions of a composite signal. However, correlation cancelers require either a secondary reference or a primary reference which correlates to either the secondary signal portion only or the primary signal portion only. For instance, for a measured signal containing noise and desirable signal, the noise can be removed with a correlation canceler if a noise reference is available. This is often the case. Although the amplitude of the reference signals are not necessarily the same as the amplitude of the corresponding primary or secondary signal portions, they have a frequency spectrum which is similar to that of the primary or secondary signal portions.

In many cases, nothing or very little is known about the secondary and/or primary signal portions. One area where measured signals comprising a primary signal portion and a secondary signal portion about which no information can easily be determined is physiological monitoring.

-2-

Physiological monitoring generally involves measured signals derived from a physiological system, such as the human body. Measurements which are typically taken with physiological monitoring systems include electrocardiographs, blood

5   pressure, blood gas saturation (such as oxygen saturation), capnographs, other blood constituent monitoring, heart rate, respiration rate, electro-encephalograph (EEG) and depth of anesthesia, for example. Other types of measurements include those which measure the pressure and quantity of a substance

10   within the body such as cardiac output, venous oxygen saturation, arterial oxygen saturation, bilirubin, total hemoglobin, breathalyzer testing, drug testing, cholesterol testing, glucose testing, extra vasation, and carbon dioxide testing, protein testing, carbon monoxide testing, and other

15   in-vivo measurements, for example. Complications arising in these measurements are often due to motion of the patient, both external and internal (muscle movement, vessel movement, and probe movement, for example), during the measurement process.

20   Many types of physiological measurements can be made by using the known properties of energy attenuation as a selected form of energy passes through a medium.

A blood gas monitor is one example of a physiological monitoring system which is based upon the measurement of

25   energy attenuated by biological tissues or substances. Blood gas monitors transmit light into the test medium and measure the attenuation of the light as a function of time. The output signal of a blood gas monitor which is sensitive to the arterial blood flow contains a component which is a

30   waveform representative of the patient's arterial pulse. This type of signal, which contains a component related to the patient's pulse, is called a plethysmographic wave, and is shown in Figure 1 as curve s. Plethysmographic waveforms are used in blood gas saturation measurements. As the heart

35   beats, the amount of blood in the arteries increases and

decreases, causing increases and decreases in energy attenuation, illustrated by the cyclic wave s in Figure 1.

Typically, a digit such as a finger, an ear lobe, or other portion of the body where blood flows close to the skin, is employed as the medium through which light energy is transmitted for blood gas attenuation measurements. The finger comprises skin, fat, bone, muscle, etc., shown schematically in Figure 2, each of which attenuates energy incident on the finger in a generally predictable and constant manner. However, when fleshy portions of the finger are compressed erratically, for example by motion of the finger, energy attenuation becomes erratic.

An example of a more realistic measured waveform S is shown in Figure 3, illustrating the effect of motion. The primary plethysmographic waveform portion of the signal s is the waveform representative of the pulse, corresponding to the sawtooth-like pattern wave in Figure 1. The large, secondary motion-induced excursions in signal amplitude obscure the primary plethysmographic signal s. Even small variations in amplitude make it difficult to distinguish the primary signal component s in the presence of a secondary signal component n.

A pulse oximeter is a type of blood gas monitor which non-invasively measures the arterial saturation of oxygen in the blood. The pumping of the heart forces freshly oxygenated blood into the arteries causing greater energy attenuation. As well understood in the art, the arterial saturation of oxygenated blood may be determined from the depth of the valleys relative to the peaks of two plethysmographic waveforms measured at separate wavelengths. Patient movement introduces motion artifacts to the composite signal as illustrated in the plethysmographic waveform illustrated in Figure 3. These motion artifacts distort the measured signal.

Summary of the Invention

This invention provides improvements upon the methods and apparatus disclosed in U.S. Patent Application No. 08/132,812, filed October 6, 1993, entitled Signal Processing Apparatus, which earlier application has been assigned to the assignee of the instant application. The present invention involves several different embodiments using the novel signal model in accordance with the present invention to isolate either a primary signal portion or a secondary signal portion of a composite measured signal. In one embodiment, a signal processor acquires a first measured signal and a second measured signal that is correlated to the first measured signal. The first signal comprises a first primary signal portion and a first secondary signal portion. The second signal comprises a second primary signal portion and a second secondary signal portion. The signals may be acquired by propagating energy through a medium and measuring an attenuated signal after transmission or reflection. Alternatively, the signals may be acquired by measuring energy generated by the medium.

In one embodiment, the first and second measured signals are processed to generate a secondary reference which does not contain the primary signal portions from either of the first or second measured signals. This secondary reference is correlated to the secondary signal portion of each of the first and second measured signals. The secondary reference is used to remove the secondary portion of each of the first and second measured signals via a correlation canceler, such as an adaptive noise canceler. The correlation canceler is a device which takes a first and second input and removes from the first input all signal components which are correlated to the second input. Any unit which performs or nearly performs this function is herein considered to be a correlation canceler.

An adaptive correlation canceler can be described by analogy to a dynamic multiple notch filter which dynamically

changes its transfer function in response to a reference signal and the measured signals to remove frequencies from the measured signals that are also present in the reference signal. Thus, a typical adaptive correlation canceler receives the signal from which it is desired to remove a component and receives a reference signal of the undesired portion. The output of the correlation canceler is a good approximation to the desired signal with the undesired component removed.

Alternatively, the first and second measured signals may be processed to generate a primary reference which does not contain the secondary signal portions from either of the first or second measured signals. The primary reference may then be used to remove the primary portion of each of the first and second measured signals via a correlation canceler. The output of the correlation canceler is a good approximation to the secondary signal with the primary signal removed and may be used for subsequent processing in the same instrument or an auxiliary instrument. In this capacity, the approximation to the secondary signal may be used as a reference signal for input to a second correlation canceler together with either the first or second measured signals for computation of, respectively, either the first or second primary signal portions.

Physiological monitors can benefit from signal processors of the present invention. Often in physiological measurements a first signal comprising a first primary portion and a first secondary portion and a second signal comprising a second primary portion and a second secondary portion are acquired. The signals may be acquired by propagating energy through a patient's body (or a material which is derived from the body, such as breath, blood, or tissue, for example) or inside a vessel and measuring an attenuated signal after transmission or reflection. Alternatively, the signal may be acquired by measuring energy generated by a patient's body, such as in

-6-

electrocardiography.   The signals are processed via the signal processor of the present invention to acquire either a secondary reference or a primary reference which is input to a correlation canceler, such as an adaptive noise canceler.

One physiological monitoring apparatus which benefits from the present invention is a monitoring system which determines a signal which is representative of the arterial pulse, called a plethysmographic wave.   This signal can be used in blood pressure calculations, blood constituent measurements, etc.   A specific example of such a use is in pulse oximetry.   Pulse oximetry involves determining the saturation of oxygen in the blood.   In this configuration, the primary portion of the signal is the arterial blood contribution to attenuation of energy as it passes through a portion of the body where blood flows close to the skin.   The pumping of the heart causes blood flow to increase and decrease in the arteries in a periodic fashion, causing periodic attenuation wherein the periodic waveform is the plethysmographic waveform representative of the arterial pulse. The secondary portion is noise.   In accordance with the present invention, the measured signals are modeled such that this secondary portion of the signal is related to the venous blood contribution to attenuation of energy as it passes through the body. The secondary portion also includes artifacts due to patient movement which causes the venous blood to flow in an unpredictable manner, causing unpredictable attenuation and corrupting the otherwise periodic plethysmographic waveform. Respiration also causes the secondary or noise portion to vary, although typically at a lower frequency than the patients pulse rate. Accordingly, the measured signal which forms a plethysmographic waveform is modeled in accordance with the present invention such that the primary portion of the signal is representative of arterial blood contribution to attenuation and the secondary portion is due to several other parameters.

-7-

A physiological monitor particularly adapted to pulse oximetry oxygen saturation measurement comprises two light emitting diodes (LED's) which emit light at different wavelengths to produce first and second signals. A detector

5   registers the attenuation of the two different energy signals after each passes through an absorptive media, for example a digit such as a finger, or an earlobe. The attenuated signals generally comprise both primary (arterial attenuator) and secondary (noise) signal portions. A static filtering

10   system, such as a bandpass filter, removes a portion of the secondary signal which is outside of a known bandwidth of interest, leaving an erratic or random secondary signal portion, often caused by motion and often difficult to remove, along with the primary signal portion.

15   A processor in accordance with one embodiment of the present invention removes the primary signal portions from the measured signals yielding a secondary reference which is a combination of the remaining secondary signal portions. The secondary reference is correlated to both of the

20   secondary signal portions. The secondary reference and at least one of the measured signals are input to a correlation canceler, such as an adaptive noise canceler, which removes the random or erratic portion of the secondary signal. This yields a good approximation to a primary plethysmographic

25   signal as measured at one of the measured signal wavelengths. As is known in the art, quantitative measurements of the amount of oxygenated arterial blood in the body can be determined from the plethysmographic signal in a variety of ways.

30   The processor of the present invention may also remove the secondary signal portions from the measured signals yielding a primary reference which is a combination of the remaining primary signal portions. The primary reference is correlated to both of the primary signal portions. The

35   primary reference and at least one of the measured signals are input to a correlation canceler which removes the primary

portions of the measured signals. This yields a good approximation to the secondary signal at one of the measured signal wavelengths. This signal may be useful for removing secondary signals from an auxiliary instrument as well as determining venous blood oxygen saturation.

In accordance with the signal model of the present invention, the two measured signals each having primary and secondary signal portions can be related by coefficients. By relating the two equations with respect to coefficients defined in accordance with the present invention, the coefficients provide information about the arterial oxygen saturation and about the noise (the venous oxygen saturation and other parameters). In accordance with this aspect of the present invention, the coefficients can be determined by minimizing the correlation between the primary and secondary signal portions as defined in the model. Accordingly, the signal model of the present invention can be utilized in many ways in order to obtain information about the measured signals as will be further apparent in the detailed description of the preferred embodiments.

One aspect of the present invention is a method for use in a signal processor in a signal processor for processing at least two measured signals S_1 and S_2 each containing a primary signal portion s and a secondary signal portion n, the signals S_1 and S_2 being in accordance with the. following relationship:

$$S_1 = s_1 + n_1$$
$$S_2 = s_2 + n_2$$

where s_1 and s_2, and n_1 and n_2 are related by:

$$s_1 = r_a s_2 \quad and \quad n_1 = r_v n_2$$

and where r_a and r_v are coefficients.

The method comprises a number of steps. A value of coefficient r_a is determined which minimize correlation between s_1 and n_1. Then, at least one of the first and second

-9-

signals is processed using the determined value for r_a to significantly reduce n from at least one of the first or second measured signal to form a clean signal.

In one embodiment, the clean signal is displayed on a display. In another embodiment, wherein the first and second signals are physiological signals, the method further comprises the step of processing the clean signal to determine a physiological parameter from the first or second measured signals. In one embodiment, the parameter is arterial oxygen saturation. In another embodiment, the parameter is an ECG signal. In yet another embodiment, wherein the first portion of the measured signals is indicative of a heart plethysmograph, the method further comprises the step of calculating the pulse rate.

Another aspect of the present invention involves a physiological monitor. The monitor has a first input configured to receive a first measured signal S_1 having a primary portion, s_1, and a secondary portion n_1. The monitor also has a second input configured to received a second measured signal S_2 having a primary portion s_2 and a secondary portion n_2. Advantageously, the first and the second measured signals S_1 and S_2 are in accordance with the following relationship:

$$S_1 = s_1 + n_1$$

$$S_2 = s_2 + n_2$$

where s_1 and s_2, and n_1 and n_2 are related by:

$$s_1 = r_a s_2 \quad and \quad n_1 = r_v n_2$$

and where r_a and r_v are coefficients.

The monitor further has a scan reference processor, the scan reference processor responds to a plurality of possible values for r_a to multiply the second measured signal by each of the possible values for r_a and for each of the resulting values, to subtract the resulting values from the first measured signal to provide a plurality of output signals. A

correlation canceler having a first input configured to receive the first measured signal, and having a second input configured to receive the plurality of output signals from the saturation scan reference processor, provides a plurality of output vectors corresponding to the correlation cancellation between the plurality of output signals and the first measured signal. An integrator having an input configured to receive the plurality of output vectors from the correlation canceler is responsive to the plurality of output vectors to determine a corresponding power for each output vector. An extremum detector is coupled at its input to the output of the integrator. The extremum detector is responsive to the corresponding power for each output vector to detect a selected power.

In one embodiment, the plurality of possible values correspond to a plurality of possible values for a selected blood constituent. In one embodiment the, the selected blood constituent is arterial blood oxygen saturation. In another embodiment, the selected blood constituent is venous blood oxygen saturation. In yet another embodiment, the selected blood constituent is carbon monoxide.

Another aspect of the present invention involves a physiological monitor. The monitor has a first input configured to receive a first measured signal S_1 having a primary portion, s_1, and a secondary portion, n_1. The monitor also has a second input configured to received a second measured signal S_2 having a primary portion s_2 and a secondary portion n_2. The first and the second measured signals S_1 and S_2 are in accordance with the following relationship:

$$S_1 = s_1 + n_1$$

$$S_2 = s_2 + n_2$$

where s_1 and s_2, and n_1 and n_2 are related by:

$$s_1 = r_a s_2 \quad and \quad n_1 = r_v n_2$$

and where r_a and r_v are coefficients.

-11-

A transform module is responsive to the first and the second measured signals and responsive to a plurality of possible values for r_a to provide at least one power curve as an output. An extremum calculation module is responsive to

5   the at least one power curve to select a value for r_a which minimizes the correlation between s and n, and to calculate from the value for r_a a corresponding saturation value as an output. A display module is responsive to the output of saturation calculation to display the saturation value.

10   Brief Description of the Drawings

Figure 1 illustrates an ideal plethysmographic waveform.

Figure 2 schematically illustrates a typical finger.

Figure 3 illustrates a plethysmographic waveform which includes a motion-induced erratic signal portion.

15   Figure 4a illustrates a schematic diagram of a physiological monitor to compute primary physiological signals.

Figure 4b illustrates a schematic diagram of a physiological monitor to compute secondary signals.

20   Figure 5a illustrates an example of an adaptive noise canceler which could be employed in a physiological monitor, to compute primary physiological signals.

Figure 5b illustrates an example of an adaptive noise canceler which could be employed in a physiological monitor,

25   to compute secondary motion artifact signals.

Figure 5c illustrates the transfer function of a multiple notch filter.

Figure 6a illustrates a schematic of absorbing material comprising N constituents within the absorbing material.

30   Figure 6b illustrates another schematic of absorbing material comprising N constituents, including one mixed layer, within the absorbing material.

Figure 6c illustrates another schematic of absorbing material comprising N constituents, including two mixed

35   layers, within the absorbing material.

Figure 7a illustrates a schematic diagram of a monitor, to compute primary and secondary signals in accordance with one aspect of the present invention.

Figure 7b illustrates the ideal correlation canceler energy or power output as a function of the signal coefficients r_1, r_2, $\cdots$ r_n. In this particular example, r_3 = r_a and r_7 = r_v.

Figure 7c illustrates the non-ideal correlation canceler energy or power output as a function of the signal coefficients r_1, r_2, $\cdots$ r_n. In this particular example, r_3 = r_a and r_7 = r_v.

Figure 8 is a schematic model of a joint process estimator comprising a least-squares lattice predictor and a regression filter.

Figure 8a is a schematic model of a joint process estimator comprising a QRD least-squares lattice (LSL) predictor and a regression filter.

Figure 9 is a flowchart representing a subroutine for implementing in software a joint process estimator as modeled in Figure 8.

Figure 9a is a flowchart representing a subroutine for implementing in software a joint process estimator as modeled in Figure 8a.

Figure 10 is a schematic model of a joint process estimator with a least-squares lattice predictor and two regression filters.

Figure 10a is a schematic model of a joint process estimator with a QRD least-squares lattice predictor and two regression filters.

Figure 11 is an example of a physiological monitor in accordance with the teachings of one aspect of the present invention.

Figure 11a illustrates an example of a low noise emitter current driver with accompanying digital to analog converter.

Figure 12 illustrates the front end analog signal conditioning circuitry and the analog to digital conversion circuitry of the physiological monitor of Figure 11.

Figure 13 illustrates further detail of the digital signal processing circuitry of Figure 11.

Figure 14 illustrates additional detail of the operations performed by the digital signal processing circuitry of Figure 11.

Figure 15 illustrates additional detail regarding the demodulation module of Figure 14.

Figure 16 illustrates additional detail regarding the decimation module of Figure 14.

Figure 17 represents a more detailed block diagram of the operations of the statistics module of Figure 14.

Figure 18 illustrates a block diagram of the operations of one embodiment of the saturation transform module of Figure 14.

Figure 19 illustrates a block diagram of the operation of the saturation calculation module of Figure 14.

Figure 20 illustrates a block diagram of the operations of the pulse rate calculation module of Figure 14.

Figure 21 illustrates a block diagram of the operations of the motion artifact suppression module of Figure 20.

Figure 21a illustrates an alternative block diagram for the operations of the motion artifact suppression module of Figure 20.

Figure 22 illustrates a saturation transform curve in accordance with the principles of the present invention.

Figure 23 illustrates a block diagram of an alternative embodiment to the saturation transform in order to obtain a saturation value.

Figure 24 illustrates a histogram saturation transform in accordance with the alternative embodiment of Figure 23.

Figures 25A-25C illustrate yet another alternative embodiment in order to obtain the saturation.

Figure 26 illustrates a signal measured at a red wavelength $\lambda a = \lambda red = 660nm$ for use in a processor of the present invention for determining the secondary reference $n'(t)$ or the primary reference $s'(t)$ and for use in a correlation canceler. The measured signal comprises a primary portion $s_{\lambda a}(t)$ and a secondary portion $n_{\lambda a}(t)$.

Figure 27 illustrates a signal measured at an infrared wavelength $\lambda b = \lambda_{IR} = 910nm$ for use in a processor of the present invention for determining the secondary reference $n'(t)$ or the primary reference $s'(t)$ and for use in a correlation canceler. The measured signal comprises a primary portion $s_{\lambda b}(t)$ and a secondary portion $n_{\lambda b}(t)$.

Figure 28 illustrates the secondary reference $n'(t)$ determined by a processor of the present invention.

Figure 29 illustrates a good approximation $s''_{\lambda a}(t)$ to the primary portion $s_{\lambda a}(t)$ of the signal $S_{\lambda a}(t)$ measured at $\lambda a = \lambda red = 660nm$ estimated by correlation cancellation with a secondary reference $n'(t)$.

Figure 30 illustrates a good approximation $s''_{\lambda b}(t)$ to the primary portion $s_{\lambda b}(t)$ of the signal $S_{\lambda b}(t)$ measured at $\lambda b = \lambda IR = 910nm$ estimated by correlation cancellation with a secondary reference $n'(t)$.

Figure 31 depicts a set of 3 concentric electrodes, i.e., a tripolar electrode sensor, to derive electrocardiography (ECG) signals, denoted as S_1, S_2 and S_3, for use with the present invention. Each of the ECG signals contains a primary portion and a secondary portion.

Detailed Description of the Invention

The present invention involves a system which utilizes first and second measured signals that each contain a primary signal portion and a secondary signal portion. In other words, given a first and second composite signals $S_1(t) = s_1(t) + n_1(t)$ and $S_2(t) = s_2(t) + n_2(t)$, the system of the present invention can be used to isolate either the primary signal portion $s(t)$ or the secondary signal portion $n(t)$. Following processing, the output of the system

-15-

provides a good approximation $n''(t)$ to the secondary signal portion $n(t)$ or a good approximation $s''(t)$ to the primary signal portion $s(t)$.

The system of the present invention is particularly useful where the primary and/or secondary signal portion $n(t)$ may contain one or more of a constant portion, a predictable portion, an erratic portion, a random portion, etc. The primary signal approximation $s''(t)$ or secondary signal approximation $n''(t)$ is derived by removing as many of the secondary signal portions $n(t)$ or primary signal portions s(t) from the composite signal $S(t)$ as possible. The remaining signal forms either the primary signal approximation $s''(t)$ or secondary signal approximation $n''(t)$, respectively. The constant portion and predictable portion of the secondary signal $n(t)$ are easily removed with traditional filtering techniques, such as simple subtraction, low pass, band pass, and high pass filtering. The erratic portion is more difficult to remove due to its unpredictable nature. If something is known about the erratic signal, even statistically, it could be removed, at least partially, from the measured signal via traditional filtering techniques. However, often no information is known about the erratic portion of the secondary signal $n(t)$. In this case, traditional filtering techniques are usually insufficient.

In order to remove the secondary signal $n(t)$, a signal model in accordance with the present invention is defined as follows for the first and second measured signals S_1 and S_2:

$$S_1 = s_1 + n_1$$

$$S_2 = s_2 + n_2$$

$$\text{with} \qquad s_1 = r_a s_2 \text{ and } n_1 = r_v n_2$$

$$\text{or} \qquad r_a = \frac{s_1}{s_2} \text{ and } n_1 = \frac{n_1}{n_2}$$

where s_1 and n_1 are at least somewhat (preferably substantially) uncorrelated and s_2 and n_2 are at least somewhat (preferably substantially) uncorrelated. The first and second measured signals S_1 and S_2 are related by correlation coefficients r_a and r_v as defined above. The use and selection of these coefficients is described in further detail below.

In accordance with one aspect of the present invention, this signal model is used in combination with a correlation canceler, such as an adaptive noise canceler, to remove or derive the erratic portion of the measured signals.

Generally, a correlation canceler has two signal inputs and one output. One of the inputs is either the secondary reference n'(t) or the primary reference s'(t) which are correlated, respectively, to the secondary signal portions n(t) and the primary signal portions s(t) present in the composite signal S(t). The other input is for the composite signal S(t). Ideally, the output of the correlation canceler s"(t) or n"(t) corresponds, respectively, to the primary signal s(t) or the secondary signal n(t) portions only. Often, the most difficult task in the application of correlation cancelers is determining the reference signals n'(t) and s'(t) which are correlated to the secondary n(t) and primary s(t) portions, respectively, of the measured signal S(t) since, as discussed above, these portions are quite difficult to isolate from the measured signal S(t). In

-17-

the signal processor of the present invention, either a secondary reference n'(t) or a primary reference s'(t) is determined from two composite signals measured simultaneously, or nearly simultaneously, at two different wavelengths, λa and λb.

A block diagram of a generic monitor incorporating a signal processor according to the present invention, and a correlation canceler is shown in Figures 4a and 4b. Two measured signals, $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$, are acquired by a detector 20. One skilled in the art will realize that for some physiological measurements, more than one detector may be advantageous. Each signal is conditioned by a signal conditioner 22a and 22b. Conditioning includes, but is not limited to, such procedures as filtering the signals to remove constant portions and amplifying the signals for ease of manipulation. The signals are then converted to digital data by an analog-to-digital converter 24a and 24b. The first measured signal $S_{\lambda a}(t)$ comprises a first primary signal portion, labeled herein $s_{\lambda a}(t)$, and a first secondary signal portion, labeled herein $n_{\lambda a}(t)$. The second measured signal $S_{\lambda b}(t)$ is at least partially correlated to the first measured signal $S_{\lambda a}(t)$ and comprises a second primary signal portion, labeled herein $s_{\lambda b}(t)$, and a second secondary signal portion, labeled herein $n_{\lambda b}(t)$. Typically the first and second secondary signal portions, $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$, are uncorrelated and/or erratic with respect to the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$. The secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ are often caused by motion of a patient in physiological measurements.

The signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ are input to a reference processor 26. The reference processor 26 multiplies the second measured signal $S_{\lambda b}(t)$ by either a factor $r_a = s_{\lambda a}(t)/s_{\lambda b}(t)$ or a factor $r_v = n_{\lambda a}(t)/n_{\lambda b}(t)$ and then subtracts the second measured signal $S_{\lambda b}(t)$ from the first measured signal $S_{\lambda a}(t)$. The signal coefficient factors r_a and r_v are determined to cause either the primary signal portions

-18-

$s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ or the secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ to cancel, respectively, when the two signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ are subtracted. Thus, the output of the reference processor 26 is either a secondary reference signal $n'(t) = n_{\lambda a}(t) - r_a n_{\lambda b}(t)$, in Figure 4a, which is correlated to both of the secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ or a primary reference signal $s'(t) = s_{\lambda a}(t) - r_v s_{\lambda b}(t)$, in Figure 4b, which is correlated to both of the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$. A reference signal $n'(t)$ or $s'(t)$ is input, along with one of the measured signals $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$, to a correlation canceler 27 which uses the reference signal $n'(t)$ or $s'(t)$ to remove either the secondary signal portions $n_{\lambda a}(t)$ or $n_{\lambda b}(t)$ or the primary signal portions $s_{\lambda a}(t)$ or $s_{\lambda b}(t)$ from the measured signal $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$. The output of the correlation canceler 27 is a good primary signal approximation $s''(t)$ or secondary signal approximation $n''(t)$. In one embodiment, the approximation $s''(t)$ or $n''(t)$ is displayed on a display 28.

In one embodiment, an adaptive noise canceler 30, an example of which is shown in block diagram form in Figure 5a, is employed as the correlation canceler 27, to remove either one of the erratic, secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ from the first and second signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$. The adaptive noise canceler 30 in Figure 5a has as one input a sample of the secondary reference $n'(t)$ which is correlated to the secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$. The secondary reference $n'(t)$ is determined from the two measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ by the processor 26 of the present invention as described herein. A second input to the adaptive noise canceler, is a sample of either the first or second composite measured signals $S_{\lambda a}(t) = s_{\lambda a}(t) + n_{\lambda a}(t)$ or $S_{\lambda b}(t) = s_{\lambda b}(t) + n_{\lambda b}(t)$.

The adaptive noise canceler 30, in Figure 5b, may also be employed to remove either one of primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ from the first and second measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$. The adaptive noise canceler 30 has as one

-19-

input a sample of the primary reference $s'(t)$ which is correlated to the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$. The primary reference $s'(t)$ is determined from the two measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ by the processor 26 of the present invention as described herein. A second input to the adaptive noise canceler 30 is a sample of either the first or second measured signals $S_{\lambda a}(t) = s_{\lambda a}(t) + n_{\lambda a}(t)$ or $S_{\lambda b}(t) = s_{\lambda b}(t) + n_{\lambda b}(t)$.

The adaptive noise canceler 30 functions to remove frequencies common to both the reference $n'(t)$ or $s'(t)$ and the measured signal $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$. Since the reference signals are correlated to either the secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ or the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$, the reference signals will be correspondingly erratic or well behaved. The adaptive noise canceler 30 acts in a manner which may be analogized to a dynamic multiple notch filter based on the spectral distribution of the reference signal $n'(t)$ or $s'(t)$.

Figure 5c illustrates an exemplary transfer function of a multiple notch filter. The notches, or dips in the amplitude of the transfer function, indicate frequencies which are attenuated or removed when a signal passes through the notch filter. The output of the notch filter is the composite signal having frequencies at which a notch is present removed. In the analogy to an adaptive noise canceler 30, the frequencies at which notches are present change continuously based upon the inputs to the adaptive noise canceler 30.

The adaptive noise canceler 30 (Figures 5a and 5b) produces an output signal, labeled herein as $s''_{\lambda a}(t)$, $s''_{\lambda b}(t)$, $n''_{\lambda a}(t)$ or $n''_{\lambda b}(t)$ which is fed back to an internal processor 32 within the adaptive noise canceler 30. The internal processor 32 automatically adjusts its own transfer function according to a predetermined algorithm such that the output of the internal processor 32 labeled $b_{\lambda}(t)$ in Figure 5a and $c_{\lambda}(t)$ in Figure 5b, closely resembles either the

-20-

secondary signal portion $n_{\lambda a}(t)$ or $n_{\lambda b}(t)$ or the primary signal portion $s_{\lambda a}(t)$ or $s_{\lambda b}(t)$. The output $b_{\lambda}(t)$ of the internal processor 32 in Figure 5a is subtracted from the measured signal, $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$, yielding a signal output $s"_{\lambda a}(t) = s_{\lambda a}(t) + n_{\lambda a}(t) - b_{\lambda a}(t)$ or a signal output $s"_{\lambda b}(t) = s_{\lambda b}(t) + n_{\lambda b}(t) - b_{\lambda b}(t)$. The internal processor optimizes $s"_{\lambda a}(t)$ or $s"_{\lambda b}(t)$ such that $s"_{\lambda a}(t)$ or $s"_{\lambda b}(t)$ is approximately equal to the primary signal $s_{\lambda a}(t)$ or $s_{\lambda b}(t)$, respectively. The output $c_{\lambda}(t)$ of the internal processor 32 in Figure 5b is subtracted from the measured signal, $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$, yielding a signal output given by $n"_{\lambda a}(t) = s_{\lambda a}(t) + n_{\lambda a}(t) - c_{\lambda a}(t)$ or a signal output given by $n"_{\lambda b}(t) = s_{\lambda b}(t) + n_{\lambda b}(t) - c_{\lambda b}(t)$. The internal processor optimizes $n"_{\lambda a}(t)$ or $n"_{\lambda b}(t)$ such that $n"_{\lambda a}(t)$ or $n"_{\lambda b}(t)$ is approximately equal to the secondary signal portion $n_{\lambda a}(t)$ or $n_{\lambda b}(t)$, respectively.

One algorithm which may be used for the adjustment of the transfer function of the internal processor 32 is a least-squares algorithm, as described in Chapter 6 and Chapter 12 of the book Adaptive Signal Processing by Bernard Widrow and Samuel Stearns, published by Prentice Hall, copyright 1985. This entire book, including Chapters 6 and 12, is hereby incorporated herein by reference.

Adaptive processors 30 in Figures 5a and 5b have been successfully applied to a number of problems including antenna sidelobe canceling, pattern recognition, the elimination of periodic interference in general, and the elimination of echoes on long distance telephone transmission lines. However, considerable ingenuity is often required to find a suitable reference signal $n'(t)$ or $s'(t)$ since the portions $n_{\lambda a}(t)$, $n_{\lambda b}(t)$, $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ cannot easily be separated from the measured composite signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$. If either the actual secondary portion $n_{\lambda a}(t)$ or $n_{\lambda b}(t)$ or the primary signal portion $s_{\lambda a}(t)$ or $s_{\lambda b}(t)$ were a priori available, techniques such as correlation cancellation would not be necessary.

An explanation which describes how the reference signals n'(t) and s'(t) may be determined follows. A first signal is measured at, for example, a wavelength λa, by a detector yielding a signal $S_{\lambda a}(t)$:

$$S_{\lambda a}(t) = s_{\lambda a}(t) + n_{\lambda a}(t) \tag{1}$$

where $s_{\lambda a}(t)$ is the primary signal portion and $n_{\lambda a}(t)$ is the secondary signal portion.

A similar measurement is taken simultaneously, or nearly simultaneously, at a different wavelength, λb, yielding:

$$S_{\lambda b}(t) = s_{\lambda b}(t) + n_{\lambda b}(t) . \tag{2}$$

Note that as long as the measurements, $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$, are taken substantially simultaneously, the secondary signal components, $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$, are correlated because any random or erratic functions affect each measurement in nearly the same fashion. The substantially predictable primary signal components, $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$, are also correlated to one another.

To obtain the reference signals n'(t) and s'(t), the measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ are transformed to eliminate, respectively, the primary or secondary signal components. In accordance with the present invention one way of doing this is to find proportionality constants, r_a and r_v, between the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ and the secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ such that the signals can be modeled as follows:

$$s_{\lambda a}(t) = r_a s_{\lambda b}(t)$$

$$\tag{3}$$

$$n_{\lambda a}(t) = r_v n_{\lambda b}(t) .$$

In accordance with the inventive signal model of the present invention, these proportionality relationships can be

satisfied in many measurements, including but not limited to absorption measurements and physiological measurements. Additionally, in accordance with the signal model of the present invention, in most measurements, the proportionality constants r_a and r_v can be determined such that:

$$n_{\lambda a}(t) \neq r_a\, n_{\lambda b}(t)$$

$$\text{(4)}$$

$$s_{\lambda a}(t) \neq r_v\, s_{\lambda b}(t).$$

Multiplying equation (2) by r_a and then subtracting equation (2) from equation (1) results in a single equation wherein the primary signal terms $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ cancel:

$$n'(t) = S_{\lambda a}(t) - r_a S_{\lambda b}(t) = n_{\lambda a}(t) - r_a n_{\lambda b}(t); \qquad \text{(5a)}$$

a non-zero signal which is correlated to each secondary signal portion $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ and can be used as the secondary reference $n'(t)$ in a correlation canceler such as an adaptive noise canceler.

Multiplying equation (2) by r_v and then subtracting equation (2) from equation (1) results in a single equation wherein the secondary signal terms $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ cancel, leaving:

$$s'(t) = S_{\lambda a}(t) - r_v S_{\lambda b}(t) = s_{\lambda a}(t) - r_v s_{\lambda b}(t); \qquad \text{(5b)}$$

a non-zero signal which is correlated to each of the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ and can be used as the signal reference $s'(t)$ in a correlation canceler such as an adaptive noise canceler.

EXAMPLE OF DETERMINATION OF PRIMARY AND SECONDARY
REFERENCE SIGNALS IN AN ABSORPTIVE SYSTEM

Correlation canceling is particularly useful in a large number of measurements generally described as absorption measurements. An example of an absorption type monitor which can advantageously employ correlation canceling, such as adaptive noise canceling, based upon a reference $n'(t)$ or

-23-

s'(t) determined by a processor of the present invention is one which determines the concentration of an energy absorbing constituent within an absorbing material when the material is subject to change. Such changes can be caused by forces about which information is desired or primary, or alternatively, by random or erratic secondary forces such as a mechanical force on the material. Random or erratic interference, such as motion, generates secondary components in the measured signal. These secondary components can be removed or derived by the correlation canceler if a suitable secondary reference n'(t) or primary reference s'(t) is known.

A schematic N constituent absorbing material comprising a container 42 having N different absorbing constituents labeled A_1, A_2, A_3, ... A_N, is shown in Figure 6a. The constituents A_1 through A_N in Figure 6a are arranged in a generally orderly, layered fashion within the container 42. An example of a particular type of absorptive system is one in which light energy passes through the container 42 and is absorbed according to the generalized Beer-Lambert Law of light absorption. For light of wavelength λa, this attenuation may be approximated by:

$$I = I_o \exp\left(-\sum_{i=1}^{N} \epsilon_{i,\lambda a} c_i x_i\right) \qquad (6)$$

Initially transforming the signal by taking the natural logarithm of both sides and manipulating terms, the signal is transformed such that the signal components are combined by addition rather than multiplication, i.e.:

$$S_{\lambda a} = \ln(I_0/I) = \sum_{i=1}^{N} \epsilon_{i,\lambda a} c_i x_i \qquad (7)$$

-24-

where I_0 is the incident light energy intensity; I is the transmitted light energy intensity; $\epsilon_{i,\lambda a}$ is the absorption coefficient of the ith constituent at the wavelength λa; $x_i(t)$ is the optical path length of ith layer, i.e., the thickness of material of the ith layer through which optical energy passes; and $c_i(t)$ is the concentration of the ith constituent in the volume associated with the thickness $x_i(t)$. The absorption coefficients ϵ_1 through ϵ_N are known values which are constant at each wavelength. Most concentrations $c_1(t)$ through $c_N(t)$ are typically unknown, as are most of the optical path lengths $x_i(t)$ of each layer. The total optical path length is the sum of each of the individual optical path lengths $x_i(t)$ of each layer.

When the material is not subject to any forces which cause change in the thicknesses of the layers, the optical path length of each layer, $x_i(t)$, is generally constant. This results in generally constant attenuation of the optical energy and thus, a generally constant offset in the measured signal. Typically, this offset portion of the signal is of little interest since knowledge about a force which perturbs the material is usually desired. Any signal portion outside of a known bandwidth of interest, including the constant undesired signal portion resulting from the generally constant absorption of the constituents when not subject to change, is removed. This is easily accomplished by traditional band pass filtering techniques. However, when the material is subject to forces, each layer of constituents may be affected by the perturbation differently than other layers. Some perturbations of the optical path lengths of each layer $x_i(t)$ may result in excursions in the measured signal which represent desired or primary information. Other perturbations of the optical path length of each layer $x_i(t)$ cause undesired or secondary excursions which mask primary information in the measured signal. Secondary signal components associated with secondary excursions must also be removed to obtain primary information from the measured

-25-

signal. Similarly, the ability to compute secondary signal components caused by secondary excursions directly allows one to obtain primary signal components from the measured signal via simple subtraction, or correlation cancellation techniques.

The correlation canceler may selectively remove from the composite signal, measured after being transmitted through or reflected from the absorbing material, either the secondary or the primary signal components caused by forces which perturb or change the material differently from the forces which perturbed or changed the material to cause respectively, either the primary or secondary signal component. For the purposes of illustration, it will be assumed that the portion of the measured signal which is deemed to be the primary signal $s_{\lambda a}(t)$ is the attenuation term $\epsilon_5 c_5 x_5(t)$ associated with a constituent of interest, namely A_5, and that the layer of constituent A_5 is affected by perturbations different than each of the layers of other constituents A_1 through A_4 and A_6 through A_N. An example of such a situation is when layer A_5 is subject to forces about which information is deemed to be primary and, additionally, the entire material is subject to forces which affect each of the layers. In this case, since the total force affecting the layer of constituent A_5 is different than the total forces affecting each of the other layers and information is deemed to be primary about the forces and resultant perturbation of the layer of constituent A_5, attenuation terms due to constituents A_1 through A_4 and A_6 through A_N make up the secondary signal portion $n_{\lambda a}(t)$. Even if the additional forces which affect the entire material cause the same perturbation in each layer, including the layer of A_5, the total forces on the layer of constituent A_5 cause it to have different total perturbation than each of the other layers of constituents A_1 through A_4 and A_6 through A_N.

It is often the case that the total perturbation affecting the layers associated with the secondary signal

-26-

components is caused by random or erratic forces. This causes the thickness of layers to change erratically and the optical path length of each layer, $x_i(t)$, to change erratically, thereby producing a random or erratic secondary signal component $n_{\lambda a}(t)$. However, regardless of whether or not the secondary signal portion $n_{\lambda a}(t)$ is erratic, the secondary signal component $n_{\lambda a}(t)$ can be either removed or derived via a correlation canceler, such as an adaptive noise canceler, having as one input, respectively, a secondary reference $n'(t)$ or a primary reference $s'(t)$ determined by a processor of the present invention as long as the perturbation on layers other than the layer of constituent A_5 is different than the perturbation on the layer of constituent A_5. The correlation canceler yields a good approximation to either the primary signal $s_{\lambda a}(t)$ or the secondary signal $n_{\lambda a}(t)$. In the event that an approximation to the primary signal is obtained, the concentration of the constituent of interest, $c_5(t)$, can often be determined since in some physiological measurements, the thickness of the primary signal component, $x_5(t)$ in this example, is known or can be determined.

The correlation canceler utilizes either the secondary reference $n'(t)$ or the primary reference $s'(t)$ determined from two substantially simultaneously measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$. $S_{\lambda a}(t)$ is determined as above in equation (7). $S_{\lambda b}(t)$ is determined similarly at a different wavelength λb. To find either the secondary reference $n'(t)$ or the primary reference $s'(t)$, attenuated transmitted energy is measured at the two different wavelengths λa and λb and transformed via logarithmic conversion. The signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ can then be written (logarithm converted) as:

$$S_{\lambda a}(t) = \epsilon_{s,\lambda a} c_5 x_5(t) + \sum_{i=1}^{4} \epsilon_{i,\lambda a} c_i x_i + \sum_{i=6}^{N} \epsilon_{i,\lambda a} c_i x_i \qquad (8)$$

$$S_{\lambda a}(t) = \epsilon_{s,\lambda a} c_5 x_5(t) + n_{\lambda a}(t) \qquad (9)$$

-27-

$$S_{\lambda b(t)} = \epsilon_{s,\lambda b} c_s x_s(t) + \sum_{i=1}^{4} \epsilon_{i,\lambda b} c_i x_i + \sum_{i=6}^{N} \epsilon_{i,\lambda b} c_i x_i \qquad (10)$$

$$S_{\lambda b(t)} = \epsilon_{s,\lambda b} c_s x_s(t) + n_{\lambda b}(t) \qquad (11)$$

Further transformations of the signals are the proportionality relationships in accordance with the signal model of the present invention defining r_a and r_v similar to equation (3), which allows determination of a noise reference $n'(t)$ and a primary reference $s'(t)$. These are:

$$\epsilon_{s,\lambda a} = r_a \epsilon_{s,\lambda b} \qquad (12a)$$

$$n_{\lambda a} = r_v n_{\lambda b} \qquad (12b)$$

where

$$n_{\lambda a} \neq r_a \, n_{\lambda b} \qquad (13a)$$

$$\epsilon_{s,\lambda a} \neq r_v \, \epsilon_{s,\lambda b}. \qquad (13b)$$

It is often the case that both equations (12) and (13) can be simultaneously satisfied. Multiplying equation (11) by r_a and subtracting the result from equation (9) yields a non-zero secondary reference which is a linear sum of secondary signal components:

$$n'(t) = S_{\lambda a}(t) - r_a S_{\lambda b}(t) = n_{\lambda a}(t) - r_a n_{\lambda b}(t) \qquad (14a)$$

$$= \sum_{i=1}^{4} \epsilon_{i,\lambda a} c_i x_i(t) + \sum_{i=6}^{N} \epsilon_{i,\lambda a} c_i x_i(t) -$$

$$\sum_{i=1}^{4} r_a \epsilon_{i,\lambda b} c_i x_i(t) + \sum_{i=6}^{N} r_a \epsilon_{i,\lambda b} c_i x_i(t) \qquad (15a)$$

$$= \sum_{i=1}^{4} c_i x_i(t) \, [\, \epsilon_{i,\lambda a} - r_a \epsilon_{i,\lambda b} \,] +$$

$$\sum_{i=6}^{N} c_i x_i(t) \, [\, \epsilon_{i,\lambda a} - r_a \epsilon_{i,\lambda b} \,] \qquad (16a)$$

Multiplying equation (11) by r_v and subtracting the result from equation (9) yields a primary reference which is a linear sum of primary signal components:

5

$$s'(t) = S_{\lambda a}(t) - r_v S_{\lambda b}(t) = s_{\lambda a}(t) - r_v s_{\lambda b}(t) \qquad (14b)$$

$$= c_s x_s(t) \epsilon_{s,\lambda a} - r_v c_s x_s(t) \epsilon_{s,\lambda b} \qquad (15b)$$

$$= c_s x_s(t) [\epsilon_{s,\lambda a} - r_v \epsilon_{s,\lambda b}]. \qquad (16b)$$

10

A sample of either the secondary reference $n'(t)$ or the primary reference $s'(t)$, and a sample of either measured signal $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$, are input to a correlation canceler 27, such as an adaptive noise canceler 30, an example of

15 which is shown in Figures 5a and 5b and a preferred example of which is discussed herein under the heading PREFERRED CORRELATION CANCELER USING A JOINT PROCESS ESTIMATOR IMPLEMENTATION. The correlation canceler 27 removes either the secondary portion $n_{\lambda a}(t)$ or $n_{\lambda b}(t)$, or the primary

20 portions, $s_{\lambda a}(t)$ or $s_{\lambda b}(t)$, of the measured signal yielding a good approximation to either the primary signals $s''_{\lambda a}(t) \approx \epsilon_{s,\lambda a} c_s x_s(t)$ or $s''_{\lambda b}(t) \approx \epsilon_{s,\lambda b} c_s x_s(t)$ or the secondary signals $n''_{\lambda a}(t) \approx n_{\lambda a}(t)$ or $n''_{\lambda b}(t) \approx n_{\lambda b}(t)$. In the event that the primary signals are obtained, the concentration $c_s(t)$ may

25 then be determined from the approximation to the primary signal $s''_{\lambda a}(t)$ or $s''_{\lambda b}(t)$ according to:

$$c_s(t) \approx s''_{\lambda a}(t)/\epsilon_{s,\lambda a} x_s(t) \qquad \qquad (17a).$$

or

30 $$c_s(t) \approx s''_{\lambda b}(t)/\epsilon_{s,\lambda b} x_s(t) \qquad (17b)$$

As discussed previously, the absorption coefficients are constant at each wavelength λa and λb and the thickness of the primary signal component, $x_s(t)$ in this example, is often known or can be determined as a function of time, thereby

35 allowing calculation of the concentration $c_s(t)$ of constituent A_s.

DETERMINATION OF CONCENTRATION OR SATURATION
IN A VOLUME CONTAINING MORE THAN ONE CONSTITUENT

Referring to Figure 6b, another material having N different constituents arranged in layers is shown. In this material, two constituents A_5 and A_6 are found within one layer having thickness $x_{5,6}(t) = x_5(t) + x_6(t)$, located generally randomly within the layer. This is analogous to combining the layers of constituents A_5 and A_6 in Figure 6a. A combination of layers, such as the combination of layers of constituents A_5 and A_6, is feasible when the two layers are under the same total forces which result in the same change of the optical path lengths $x_5(t)$ and $x_6(t)$ of the layers.

Often it is desirable to find the concentration or the saturation, i.e., a percent concentration, of one constituent within a given thickness which contains more than one constituent and is subject to unique forces. A determination of the concentration or the saturation of a constituent within a given volume may be made with any number of constituents in the volume subject to the same total forces and therefore under the same perturbation or change. To determine the saturation of one constituent in a volume comprising many constituents, as many measured signals as there are constituents which absorb incident light energy are necessary. It will be understood that constituents which do not absorb light energy are not consequential in the determination of saturation. To determine the concentration, as many signals as there are constituents which absorb incident light energy are necessary as well as information about the sum of concentrations.

It is often the case that a thickness under unique motion contains only two constituents. For example, it may be desirable to know the concentration or saturation of A_5 within a given volume which contains A_5 and A_6. In this case, the primary signals $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ comprise terms related to both A_5 and A_6 so that a determination of the concentration or saturation of A_5 or A_6 in the volume may be made. A

-30-

determination of saturation is discussed herein. It will be understood that the concentration of A_5 in a volume containing both A_5 and A_6 could also be determined if it is known that $A_5+A_6=1$, i.e., that there are no constituents in the volume which do not absorb incident light energy at the particular measurement wavelengths chosen. The measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ can be written (logarithm converted) as:

$$S_{\lambda a}(t) = \epsilon_{5,\lambda a}c_5 x_{5,6}(t) + \epsilon_{6,\lambda a}c_6 x_{5,6}(t) + n_{\lambda a}(t) \qquad (18a)$$

$$= s_{\lambda a}(t) + n_{\lambda a}(t) \qquad (18b)$$

$$S_{\lambda b}(t) = \epsilon_{5,\lambda b}c_5 x_{5,6}(t) + \epsilon_{6,\lambda b}c_6 x_{5,6}(t) + n_{\lambda b}(t) \qquad (19a)$$

$$= s_{\lambda b}(t) + n_{\lambda b}(t). \qquad (19b)$$

It is also often the case that there may be two or more thicknesses within a medium each containing the same two constituents but each experiencing a separate motion as ·in Figure 6c. For example, it may be desirable to know the concentration or saturation of A_5 within a given volume which contains A_5 and A_6 as well as the concentration or saturation of A_3 within a given volume which contains A_3 and A_4, A_3 and A_4 having the same constituency as A_5 and A_6 respectively. In this case, the primary signals $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ again comprise terms related to both A_5 and A_6 and portions of the secondary signals $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ comprise terms related to both A_3 and A_4. The layers, A_3 and A_4, do not enter into the primary equation because they are assumed to be perturbed by a different frequency, or random or erratic secondary forces which are uncorrelated with the primary force. Since constituents 3 and 5 as well as constituents 4 and 6 are taken to be the same, they have the same absorption coefficients (i.e., $\epsilon_{3,\lambda a} = \epsilon_{5,\lambda a}$; $\epsilon_{3,\lambda b} = \epsilon_{5,\lambda b}$; $\epsilon_{4,\lambda a} = \epsilon_{6,\lambda a}$ and $\epsilon_{4,\lambda b} = \epsilon_{6,\lambda b}$. Generally speaking, however, A_3 and A_4 will have

-31-

different concentrations than A_5 and A_6 and will therefore have a different saturation. Consequently a single constituent within a medium may have one or more saturations associated with it. The primary and secondary signals according to this model may be written as:

$$s_{\lambda a}(t) = [\epsilon_{5,\lambda a}c_5 + \epsilon_{6,\lambda a}c_6] \ x_{5,6}(t) \qquad (20a)$$

$$n_{\lambda a}(t) = [\epsilon_{5,\lambda a}c_3 + \epsilon_{6,\lambda a}c_4] \ x_{3,4}(t)$$

$$+ \ \sum_{i=1}^{2} \epsilon_{i,\lambda a} \ c_i \ x_i(t) + \sum_{i=7}^{N} \epsilon_{i,\lambda a} \ c_i \ x_i(t) \qquad (20b)$$

$$n_{\lambda a}(t) = [\epsilon_{5,\lambda a}c_3 + \epsilon_{6,\lambda a}c_4] \ x_{3,4}(t) + n_{\lambda a}(t) \qquad (20c)$$

$$s_{\lambda b}(t) = [\epsilon_{5,\lambda b}c_5 + \epsilon_{6,\lambda b}c_6] \ x_{5,6}(t) \qquad (21a)$$

$$n_{\lambda b}(t) = [\epsilon_{5,\lambda b}c_3 + \epsilon_{6,\lambda b}c_4] \ x_{3,4}(t)$$

$$+ \ \sum_{i=1}^{2} \epsilon_{i,\lambda b} \ c_i x_i(t) + \sum_{i=7}^{N} \epsilon_{i,\lambda b}c_i x_i(t) . \qquad (21b)$$

$$n_{\lambda b}(t) = [\epsilon_{5,\lambda b}c_3 + \epsilon_{6,\lambda b}c_4] \ x_{3,4}(t) + n_{\lambda b}(t) \qquad (21c)$$

where signals $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ are similar to the secondary signals $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ except for the omission of the 3, 4 layer.

Any signal portions whether primary or secondary, outside of a known bandwidth of interest, including the constant undesired secondary signal portion resulting from the generally constant absorption of the constituents when not under perturbation, should be removed to determine an approximation to either the primary signal or the secondary signal within the bandwidth of interest. This is easily

accomplished by traditional band pass filtering techniques. As in the previous example, it is often the case that the total perturbation or change affecting the layers associated with the secondary signal components is caused by random or erratic forces, causing the thickness of each layer, or the optical path length of each layer, $x_i(t)$, to change erratically, producing a random or erratic secondary signal component $n_{\lambda a}(t)$. Regardless of whether or not the secondary signal portion $n_{\lambda a}(t)$ is erratic, the secondary signal component $n_{\lambda a}(t)$ can be removed or derived via a correlation canceler, such as an adaptive noise canceler, having as one input a secondary reference $n'(t)$ or a primary reference $s'(t)$ determined by a processor of the present invention as long as the perturbation in layers other than the layer or constituents A_5 and A_6 is different than the perturbation in the layer of constituents A_5 and A_6. Either the erratic secondary signal components $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ or the primary components $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ may advantageously be removed from equations (18) and (19), or alternatively equations (20) and (21), by a correlation canceler. The correlation canceler, again, requires a sample of either the primary reference $s'(t)$ or the secondary reference $n'(t)$ and a sample of either of the composite signals $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$ of equations (18) and (19).

DETERMINATION OF PRIMARY AND SECONDARY REFERENCE SIGNALS
FOR SATURATION MEASUREMENTS

One method for determining reference signals $s'(t)$ or $n'(t)$ from the measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ in accordance with one aspect of the invention is what will be referred to as the constant saturation approach. In this approach, it is assumed that the saturation of A_5 in the volume containing A_5 and A_6 and the saturation of A_3 in the volume containing A_3 and A_4 remains relatively constant over some period of time, i.e.:

$$\text{Saturation}(A_5(t)) = c_5(t) \, / \, [c_5(t) + c_6(t)] \qquad (22a)$$

-33-

$$\text{Saturation}(A_3(t)) = c_3(t) \ / \ [c_3(t) + c_4(t)] \qquad (22b)$$

$$\text{Saturation}(A_5(t)) = \{1 + [c_6(t)/c_5(t)] \}^{-1} \qquad (23a)$$

$$\text{Saturation}(A_3(t)) = \{1 + [c_4(t)/c_3(t)]\}^{-1} \qquad (23b)$$

are substantially constant over many samples of the measured signals $S_{\lambda a}$ and $S_{\lambda b}$. This assumption is accurate over many samples since saturation generally changes relatively slowly in physiological systems.

The constant saturation assumption is equivalent to assuming that:

$$c_5(t)/c_6(t) = \text{constant}_1 \qquad (24a)$$

$$c_3(t)/c_4(t) = \text{constant}_2 \qquad (24b)$$

since the only other term in equations (23a) and (23b) is a constant, namely the numeral 1.

Using this assumption, the proportionality constants r_a and r_v which allow determination of the secondary reference signal $n'(t)$ and the primary reference signal $s'(t)$ in the constant saturation method are:

$$r_a = \frac{\epsilon_{5,\lambda a}\, c_5\, x_{5,6}(t) + \epsilon_{6,\lambda a}\, c_6\, x_{5,6}(t)}{\epsilon_{5,\lambda b}\, c_5\, x_{5,6}(t) + \epsilon_{6,\lambda b}\, c_6\, x_{5,6}(t)} \qquad (25a)$$

$$= s_{\lambda a}(t)/s_{\lambda b}(t) \qquad (26a)$$

$$= \frac{\epsilon_{5,\lambda a}\, c_5 + \epsilon_{6,\lambda a} c_6}{\epsilon_{5,\lambda b} c_5 + \epsilon_{6,\lambda b} c_6} \qquad (27a)$$

$$= \frac{\epsilon_{5,\lambda a}\, (c_5/c_6) + \epsilon_{6,\lambda a}}{\epsilon_{5,\lambda b}\, (c_5/c_6) + \epsilon_{6,\lambda b}} \qquad (28a)$$

-34-

$$\sim \quad s''_{\lambda a}(t)/s''_{\lambda b}(t) = \text{constant}_3; \text{ where} \qquad (29a)$$

$$n_{\lambda a}(t) \neq r_a(t)n_{\lambda b}(t) \qquad (30a)$$

and

$$r_v = \frac{\epsilon_{5,\lambda a} \, c_3 \, x_{3,4}(t) + \epsilon_{6,\lambda a} \, c_4 \, x_{3,4}(t)}{\epsilon_{5,\lambda b}c_3 x_{3,4}(t) + \epsilon_{6,\lambda b}c_4 x_{3,4}(t))} \qquad (25b)$$

$$= n_{\lambda a}(t)/n_{\lambda b}(t) \qquad (26b)$$

$$= \frac{\epsilon_{5,\lambda a}c_3 + \epsilon_{6,\lambda a}c_4}{\epsilon_{5,\lambda b}c_2 + \epsilon_{6,\lambda b}c_4} \qquad (27b)$$

$$= \frac{\epsilon_{5,\lambda a}(c_3/c_4) + \epsilon_{6,\lambda a}}{\epsilon_{5,\lambda b}(c_3/c_4) + \epsilon_{6,\lambda b}} \qquad (28b)$$

$$\sim \quad n''_{\lambda a}(t)/n''_{\lambda b}(t) = \text{constant}_4; \text{ where} \qquad (29b)$$

$$s_{\lambda a}(t) \neq r_v(t)s_{\lambda b}(t). \qquad (30b)$$

In accordance with the present invention, it is often the case that both equations (26) and (30) can be simultaneously satisfied to determine the proportionality constants r_a and r_v. Additionally, the absorption coefficients at each wavelength $\epsilon_{5,\lambda a}$, $\epsilon_{6,\lambda a}$, $\epsilon_{5,\lambda b}$, and $\epsilon_{6,\lambda b}$ are constant and the central assumption of the constant saturation method is that $c_5(t)/c_6(t)$ and $c_3(t)/c_4(t)$ are constant over many sample periods. Thus, new proportionality constants r_a and r_v may be determined every few samples from new approximations to either the primary or secondary signal as output from the correlation canceler. Thus, the approximations to either the primary signals $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ or the secondary signals $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$, found by the correlation canceler for a substantially immediately

preceding set of samples of the measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ are used in a processor of the present invention for calculating the proportionality constants, r_a and r_v, for the next set of samples of the measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$.

Multiplying equation (19) by r_a and subtracting the resulting equation from equation (18) yields a non-zero secondary reference signal:

$$n'(t) = S_{\lambda a}(t) - r_a S_{\lambda b}(t) = n_{\lambda a}(t) - r_a n_{\lambda b}(t). \qquad (31a)$$

Multiplying equation (19) by r_v and subtracting the resulting equation from equation (18) yields a non-zero primary reference signal:

$$s'(t) = S_{\lambda a}(t) - r_v S_{\lambda b}(t) = s_{\lambda a}(t) - r_v s_{\lambda b}(t). \qquad (31b)$$

When using the constant saturation method in patient monitoring, initial proportionality coefficients can be determined as further explained below. It is not necessary for the patient to remain motionless even for an initialization period. With values for the proportionality coefficients r_a and r_v determined, a correlation canceler may be utilized with a secondary reference $n'(t)$ or a primary reference $s'(t)$.

DETERMINATION OF SIGNAL COEFFICIENTS FOR
PRIMARY AND SECONDARY REFERENCE SIGNALS
USING THE CONSTANT SATURATION METHOD

In accordance with one aspect of the present invention, the reference processor 26 of Figures 4a and Figure 4b of the present invention may be configured to multiply the second measured assumed signal $S_{\lambda b}(t) = s_{\lambda b}(t) + n_{\lambda b}(t)$ by each of a plurality of signal coefficients $r_1, r_2, \cdots r_n$ and then subtract each result from the first measured signal $S_{\lambda a}(t) = s_{\lambda a}(t) + n_{\lambda a}(t)$ to obtain a plurality of reference signals

$$R'(r, t) = s_{\lambda a}(t) - r\, s_{\lambda b}(t) + n_{\lambda a}(t) - r\, n_{\lambda b}(t) \qquad (32)$$

-36-

for $r = r_1$, r_2, $\cdots$ r_n as shown in Figure 7a. In other words, a plurality of signal coefficients are chosen to represent a cross section of possible signal coefficients.

In order to determine either the primary reference $s'(t)$ or the secondary reference $n'(t)$ from the above plurality of reference signals of equation (32), signal coefficients r_a and r_v are determined from the plurality of assumed signal coefficients r_1, r_2, $\cdots$ r_n. The coefficients r_a and r_v are selected such that they cause either the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ or the secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ to cancel or nearly cancel when they are substituted into the reference function $R'(r, t)$, e. g.

$$s_{\lambda a}(t) = r_a s_{\lambda b}(t) \tag{33a}$$

$$n_{\lambda a}(t) = r_v n_{\lambda b}(t) \tag{33b}$$

$$n'(t) = R'(r_a, t) = n_{\lambda a}(t) - r_a n_{\lambda b}(t) \tag{33c}$$

$$s'(t) = R'(r_v, t) = s_{\lambda a}(t) - r_v s_{\lambda b}(t). \tag{33d}$$

In other words, coefficients r_a and r_v are selected at values which reflect the minimum of correlation between the primary signal portions and the secondary signal portions. In practice, one does not usually have significant prior information about either the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ or the secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ of the measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$. The lack of this information makes it difficult to determine which of the plurality of coefficients r_1, r_2, $\cdots$ r_n correspond to the signal coefficients $r_a = s_{\lambda a}(t)/s_{\lambda b}(t)$ and $r_v = n_{\lambda a}(t)/n_{\lambda b}(t)$.

One approach to determine the signal coefficients r_a and r_v from the plurality of coefficients r_1, r_2, $\cdots$ r_n employs the use of a correlation canceler 27, such as an adaptive noise canceler, which takes a first input which corresponds to one of the measured signals $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$ and takes a

-37-

second input which corresponds to successively each one of the plurality of reference signals $R'(r_1, t)$, $R'(r_2, t)$, $\cdots$, $R'(r_n, t)$ as shown in Figure 7a. For each of the reference signals $R'(r_1, t)$, $R'(r_2, t)$, $\cdots$, $R'(r_n, t)$ the corresponding

5 output of the correlation canceler 27 is input to a "squares" operation 28 which squares the output of the correlation canceler 27. The output of the squares operation 28 is provided to an integrator 29 for forming a cumulative output signal (a summation of the squares). The cumulative output

10 signal is subsequently input to an extremum detector 31. The purpose of the extremum detector 31 is to chose signal coefficients r_a and r_v from the set r_1, r_2, $\cdots$ r_n by observing which provide a maximum in the cumulative output signal as in Figures 7b and 7c. In other words, coefficients which

15 provide a maximum integrated output, such as energy or power, from the correlation-canceler 27 correspond to the signal coefficients r_a and r_v which relate to a minimum correlation between the primary signal portions and the secondary signal portions in accordance with the signal model of the present

20 invention. One could also configure a system geometry which would require one to locate the coefficients from the set r_1, r_2, $\cdots$ r_n which provide a minimum or inflection in the cumulative output signal to identify the signal coefficients r_a and r_v.

25 Use of a plurality of coefficients in the processor of the present invention in conjunction with a correlation canceler 27 to determine the signal coefficients r_a and r_v may be demonstrated by using the properties of correlation cancellation. If x, y and z are taken to be any collection

30 of three time varying signals, then the properties of some correlation cancelers $C(x, y)$ may be defined as follows:

Property (1) $C(x, y) = 0$ for x, y correlated (34a)
Property (2) $C(x, y) = x$ for x, y uncorrelated (34b)

35 Property (3) $C(x + y, z) = C(x, z) + C(y, z)$ (34c)

With properties (1), (2) and (3) it is easy to demonstrate that the energy or power output of a correlation canceler with a first input which corresponds to one of the measured signals $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$ and a second input which corresponds

5   to successively each one of a plurality of reference signals $R'(r_1, t)$, $R'(r_2, t)$, $\cdots$, $R'(r_n, t)$ can determine the signal coefficients r_a and r_v needed to produce the primary reference $s'(t)$ and secondary reference $n'(t)$. If we take as a first input to the correlation canceler the measured signal $S_{\lambda a}(t)$

10   and as a second input the plurality of reference signals $R'(r_1, t)$, $R'(r_2, t)$, $\cdots$, $R'(r_n, t)$ then the outputs of the correlation canceler $C(S_{\lambda a}(t), R'(r_j, t))$ for $j = 1, 2, \cdots, n$ may be written as

15     $C(s_{\lambda a}(t) + n_{\lambda a}(t), s_{\lambda a}(t) - r_j s_{\lambda b}(t) + n_{\lambda a}(t) - r_j n_{\lambda b}(t))$ (35)

where $j = 1, 2, \cdots, n$ and we have used the expressions

$$R'(r, t) = S_{\lambda a}(t) - r S_{\lambda b}(t) \qquad (36)$$

20

$$S_{\lambda a}(t) = s_{\lambda a}(t) + n_{\lambda a}(t) \qquad (37a)$$

$$S_{\lambda b}(t) = s_{\lambda b}(t) + n_{\lambda b}(t). \qquad (37b)$$

25   The use of property (3) allows one to expand equation (35) into two terms

$$C(S_{\lambda a}(t), R'(r,t)) = C(s_{\lambda a}(t), s_{\lambda a}(t) - r s_{\lambda b}(t) + n_{\lambda a}(t) - r n_{\lambda b}(t))$$

30     $$+ C(n_{\lambda a}(t), s_{\lambda a}(t) - r s_{\lambda b}(t) + n_{\lambda a}(t) - r n_{\lambda b}(t)) \qquad (38)$$

so that upon use of properties (1) and (2) the correlation canceler output is given by

$$C(S_{\lambda a}(t), R'(r_j, t)) = s_{\lambda a}(t) \delta(r_j - r_a) + n_{\lambda a}(t) \delta(r_j - r_v) \qquad (39)$$

35

where $\delta(x)$ is the unit impulse function

-39-

$$\delta(x) = 0 \text{ if } x \neq 0$$

$$\delta(x) = 1 \text{ if } x = 0. \tag{40}$$

The time variable, t, of the correlation canceler output $C(S_{\lambda a}(t), R'(r_j, t))$ may be eliminated by computing its energy or power. The energy of the correlation canceler output is given by

$$E_{\lambda a}(r_j) = \int C^2(S_{\lambda a}(t), R'(r_j, t)) dt$$
$$= \delta(r_j - r_a) \int s^2_{\lambda a}(t) dt + \delta(r_j - r_v) \int n^2_{\lambda a}(t) dt. \tag{41a}$$

It should be understood that one could, equally well, have chosen the measured signal $S_{\lambda b}(t)$ as the first input to the correlation canceler and the plurality of reference signals $R'(r_1, t)$, $R'(r_2, t)$, $\cdots$, $R'(r_n, t)$ as the second input. In this event, the correlation canceler energy output is

$$E_{\lambda b}(r_j) = \int C^2(S_{\lambda b}(t), R'(r, t)) dt \tag{41b}$$
$$= \delta(r_j - r_a) \int s^2_{\lambda b}(t) dt + \delta(r_j - r_v) \int n^2_{\lambda b}(t) dt.$$

It should also be understood that in practical situations the use of discrete time measurement signals may be employed as well as continuous time measurement signals. A system which performs a discrete transform (e.g., a saturation transform in the present example) in accordance with the present invention is described with reference to Figures 11-22. In the event that discrete time measurement signals are used, integration approximation methods such as the trapezoid rule, midpoint rule, Tick's rule, Simpson's approximation or other techniques may be used to compute the correlation canceler energy or power output. In the discrete time measurement signal case, the energy output of the correlation canceler may be written, using the trapezoid rule, as

$$E_{\lambda a}(r_j) = \delta(r_j - r_a) \Delta t \{ \sum_{i=0}^{n} s^2_{\lambda a}(t_i) - 0.5(s^2_{\lambda a}(t_0) + s^2_{\lambda a}(t_n)) \}$$

$$+\delta(r_j - r_v) \Delta t \{ \sum_{i=0}^{n} n^2_{\lambda a}(t_i) - 0.5(n^2_{\lambda a}(t_0) + n^2_{\lambda a}(t_n)) \} \qquad (42a)$$

$$E_{\lambda b}(r) = \delta(r_j - r_a) \Delta t \{ \sum_{i=0}^{n} s^2_{\lambda b}(t_i) - 0.5(s^2_{\lambda b}(t_0) + s^2_{\lambda b}(t_n)) \}$$

$$+\delta(r_j - r_v) \Delta t \{ \sum_{i=0}^{n} n^2_{\lambda b}(t_i) - 0.5(n^2_{\lambda b}(t_0) + n^2_{\lambda b}(t_n)) \} \qquad (42b)$$

where t_i is the ith discrete time, t_0 is the initial time, t_n is the final time and Δt is the time between discrete time measurement samples.

The energy functions given above, and shown in Figure 7b, indicate that the correlation canceler output is usually zero due to correlation between the measured signal $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$ and many of the plurality of reference signals $R'(r_1, t)$, $R'(r_2, t)$, $\cdots$, $R'(r_n, t)$. However, the energy functions are non zero at values of r_j which correspond to cancellation of either the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ or the secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ in the reference signal $R'(r_j, t)$. These values correspond to the signal coefficients r_a and r_v.

It should be understood that there may be instances in time when either the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ or the secondary signal portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ are identically zero or nearly zero. In these cases, only one signal coefficient value will provide maximum energy or power output of the correlation canceler.

Since there may be more than one signal coefficient value which provides maximum correlation canceler energy or power output, an ambiguity may arise. It may not be immediately obvious which signal coefficient together with the reference function $R'(r, t)$ provides either the primary or secondary reference. In such cases, it is necessary to consider the constraints of the physical system at hand. For example, in pulse oximetry, it is known that arterial blood, whose signature is the primary plethysmographic wave, has greater oxygen saturation than venous blood, whose signature is the secondary erratic or random signal. Consequently, in pulse oximetry, the ratio of the primary signals due to arterial pulsation $r_a = s_{\lambda a}(t)/s_{\lambda b}(t)$ is the smaller of the two signal coefficient values while the ratio of the secondary signals due to mainly venous blood dynamics $r_v = n_{\lambda a}(t)/n_{\lambda b}(t)$ is the larger of the two signal coefficient values, assuming $\lambda a = 660$ nm and $\lambda b = 910$ nm.

It should also be understood that in practical implementations of the plurality of reference signals and cross correlator technique, the ideal features listed as properties (1), (2) and (3) above will not be precisely satisfied but will be approximations thereof. Therefore, in practical implementations of this embodiment of the present invention, the correlation canceler energy curves depicted in Figure 7b will not consist of infinitely narrow delta functions but will have finite width associated with them as depicted in Figure 7c.

It should also be understood that it is possible to have more than two signal coefficient values which produce maximum energy or power output from a correlation canceler. This situation arises when the measured signals each contain more than two components each of which are related by a ratio as follows:

-42-

$$S_{\lambda a}(t) = \sum_{i=0}^{n} f_{\lambda a, i}(t)$$

(43)

$$S_{\lambda b}(t) = \sum_{i=0}^{n} f_{\lambda b, i}(t)$$

5   where

$$f_{\lambda a, i}(t) = r_i \, f_{\lambda b, i}(t) \qquad i = 1, \ldots, n$$
$$r_i \neq r_j.$$

Thus, reference signal techniques together with a
10 correlation cancellation, such as an adaptive noise canceler,
can be employed to decompose a signal into two or more signal
components each of which is related by a ratio.

PREFERRED CORRELATION CANCELER USING A
JOINT PROCESS ESTIMATOR IMPLEMENTATION

15   Once either the secondary reference n'(t) or the primary
reference s'(t) is determined by the processor of the present
invention, the correlation canceler can be implemented in
either hardware or software. The preferred implementation of
a correlation canceler is that of an adaptive noise canceler
20 using a joint process estimator.

The least mean squares (LMS) implementation of the
internal processor 32 described above in conjunction with the
adaptive noise canceler of Figures 5a and Figure 5b is
relatively easy to implement, but lacks ·the speed of
25 adaptation desirable for most physiological monitoring
applications of the present invention. Thus, a faster
approach for adaptive noise canceling, called a least-squares
lattice joint process estimator model, is used in one
embodiment. A joint process estimator 60 is shown
30 diagrammatically in Figure 8 and is described in detail in
Chapter 9 of <u>Adaptive Filter Theory</u> by Simon Haykin,

published by Prentice-Hall, copyright 1986. This entire book, including Chapter 9, is hereby incorporated herein by reference.

The function of the joint process estimator is to remove either the secondary signal portions $n_{\lambda a}(t)$ or $n_{\lambda b}(t)$ or the primary signal portions $s_{\lambda a}(t)$ or $s_{\lambda b}(t)$ from the measured signals $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$, yielding either a primary signal approximation $s''_{\lambda a}(t)$ or $s''_{\lambda b}(t)$ or a secondary signal approximation $n''_{\lambda a}(t)$ or $n''_{\lambda b}(t)$. Thus, the joint process estimator estimates either the value of the primary signals $s_{\lambda a}(t)$ or $s_{\lambda b}(t)$ or the secondary signals $n_{\lambda a}(t)$ or $n_{\lambda b}(t)$. The inputs to the joint process estimator 60 are either the secondary reference $n'(t)$ or the primary reference $s'(t)$ and the composite measured signal $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$. The output is a good approximation to the signal $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$ with either the secondary signal or the primary signal removed, i.e. a good approximation to either $s_{\lambda a}(t)$, $s_{\lambda b}(t)$, $n_{\lambda a}(t)$ or $n_{\lambda b}(t)$.

The joint process estimator 60 of Figure 8 utilizes, in conjunction, a least square lattice predictor 70 and a regression filter 80. Either the secondary reference $n'(t)$ or the primary reference $s'(t)$ is input to the least square lattice predictor 70 while the measured signal $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$ is input to the regression filter 80. For simplicity in the following description, $S_{\lambda a}(t)$ will be the measured signal from which either the primary portion $s_{\lambda a}(t)$ or the secondary portion $n_{\lambda a}(t)$ will be estimated by the joint process estimator 60. However, it will be noted that $S_{\lambda b}(t)$ could also be input to the regression filter 80 and the primary portion $s_{\lambda b}(t)$ or the secondary portion $n_{\lambda b}(t)$ of this signal could be estimated.

The joint process estimator 60 removes all frequencies that are present in both the reference $n'(t)$ or $s'(t)$, and the measured signal $S_{\lambda a}(t)$. The secondary signal portion $n_{\lambda a}(t)$ usually comprises frequencies unrelated to those of the primary signal portion $s_{\lambda a}(t)$. It is improbable that the

secondary signal portion $n_{\lambda a}(t)$ would be of exactly the same spectral content as the primary signal portion $s_{\lambda a}(t)$. However, in the unlikely event that the spectral content of $s_{\lambda a}(t)$ and $n_{\lambda a}(t)$ are similar, this approach will not yield

5   accurate results. Functionally, the joint process estimator 60 compares the reference input signal $n'(t)$ or $s'(t)$, which is correlated to either the secondary signal portion $n_{\lambda a}(t)$ or the primary signal portion $s_{\lambda a}(t)$, and input signal $S_{\lambda a}(t)$ and removes all frequencies which are identical. Thus, the

10 joint process estimator 60 acts as a dynamic multiple notch filter to remove those frequencies in the secondary signal component $n_{\lambda a}(t)$ as they change erratically with the motion of the patient or those frequencies in the primary signal component $s_{\lambda a}(t)$ as they change with the arterial pulsation

15 of the patient. This yields a signal having substantially the same spectral content and amplitude as either the primary signal $s_{\lambda a}(t)$ or the secondary signal $n_{\lambda a}(t)$. Thus, the output $s''_{\lambda a}(t)$ or $n''_{\lambda a}(t)$ of the joint process estimator 60 is a very good approximation to either the primary signal $s_{\lambda a}(t)$

20 or the secondary signal $n_{\lambda a}(t)$.

The joint process estimator 60 can be divided into stages, beginning with a zero-stage and terminating in an mth-stage, as shown in Figure 8. Each stage, except for the zero-stage, is identical to every other stage. The zero-

25 stage is an input stage for the joint process estimator 60. The first stage through the mth-stage work on the signal produced in the immediately previous stage, i.e., the $(m-1)^{th}$-stage, such that a good primary signal approximation $s''_{\lambda a}(t)$ or secondary signal approximation $n''_{\lambda a}(t)$ is produced as

30 output from the mth-stage.

The least-squares lattice predictor 70 comprises registers 90 and 92, summing elements 100 and 102, and delay elements 110. The registers 90 and 92 contain multiplicative values of a forward reflection coefficient $\Gamma_{f,m}(t)$ and a

35 backward reflection coefficient $\Gamma_{b,m}(t)$ which multiply the reference signal $n'(t)$ or $s'(t)$ and signals derived from the

-45-

reference signal $n'(t)$ or $s'(t)$. Each stage of the least-squares lattice predictor outputs a forward prediction error $f_m(t)$ and a backward prediction error $b_m(t)$. The subscript m is indicative of the stage.

For each set of samples, i.e. one sample of the reference signal $n'(t)$ or $s'(t)$ derived substantially simultaneously with one sample of the measured signal $S_{\lambda a}(t)$, the sample of the reference signal $n'(t)$ or $s'(t)$ is input to the least-squares lattice predictor 70. The zero-stage forward prediction error $f_0(t)$ and the zero-stage backward prediction error $b_0(t)$ are set equal to the reference signal $n'(t)$ or $s'(t)$. The backward prediction error $b_0(t)$ is delayed by one sample period by the delay element 110 in the first stage of the least-squares lattice predictor 70. Thus, the immediately previous value of the reference $n'(t)$ or $s'(t)$ is used in calculations involving the first-stage delay element 110. The zero-stage forward prediction error is added to the negative of the delayed zero-stage backward prediction error $b_0(t-1)$ multiplied by the forward reflection coefficient value $\Gamma_{f,1}(t)$ register 90 value, to produce a first-stage forward prediction error $f_1(t)$. Additionally, the zero-stage forward prediction error $f_0(t)$ is multiplied by the backward reflection coefficient $\Gamma_{b,1}(t)$ register 92 value and added to the delayed zero-stage backward prediction error $b_0(t-1)$ to produce a first-stage backward prediction error $b_1(t)$. In each subsequent stage, m, of the least square lattice predictor 70, the previous forward and backward prediction error values, $f_{m-1}(t)$ and $b_{m-1}(t-1)$, the backward prediction error being delayed by one sample period, are used to produce values of the forward and backward prediction errors for the present stage, $f_m(t)$ and $b_m(t)$.

The backward prediction error $b_m(t)$ is fed to the concurrent stage, m, of the regression filter 80. There it is input to a register 96, which contains a multiplicative regression coefficient value $\kappa_{m,\lambda a}(t)$. For example, in the

zero-stage of the regression filter 80, the zero-stage backward prediction error $b_0(t)$ is multiplied by the zero-stage regression coefficient $\kappa_{0,\lambda a}(t)$ register 96 value and subtracted from the measured value of the signal $S_{\lambda a}(t)$ at a summing element 106 to produce a first stage estimation error signal $e_{1,\lambda a}(t)$. The first-stage estimation error signal $e_{1,\lambda a}(t)$ is a first approximation to either the primary signal or the secondary signal. This first-stage estimation error signal $e_{1,\lambda a}(t)$ is input to the first-stage of the regression filter 80. The first-stage backward prediction error $b_1(t)$, multiplied by the first-stage regression coefficient $\kappa_{1,\lambda a}(t)$ register 96 value is subtracted from the first-stage estimation error signal $e_{1,\lambda a}(t)$ to produce the second-stage estimation error $e_{2,\lambda a}(t)$. The second-stage estimation error signal $e_{2,\lambda a}(t)$ is a second, somewhat better approximation to either the primary signal $s_{\lambda a}(t)$ or the secondary signal $n_{\lambda a}(t)$.

The same processes are repeated in the least-squares lattice predictor 70 and the regression filter 80 for each stage until a good approximation $e_{m,\lambda a}(t)$, to either the primary signal $s_{\lambda a}(t)$ or the secondary signal $n_{\lambda a}(t)$ is determined. Each of the signals discussed above, including the forward prediction error $f_m(t)$, the backward prediction error $b_m(t)$, the estimation error signal $e_{m,\lambda a}(t)$, is necessary to calculate the forward reflection coefficient $\Gamma_{f,m}(t)$, the backward reflection coefficient $\Gamma_{b,m}(t)$, and the regression coefficient $\kappa_{m,\lambda a}(t)$ register 90, 92, and 96 values in each stage, m. In addition to the forward prediction error $f_m(t)$, the backward prediction error $b_m(t)$, and the estimation error $e_{m,\lambda a}(t)$ signals, a number of intermediate variables, not shown in Figure 8 but based on the values labeled in Figure 8, are required to calculate the forward reflection coefficient $\Gamma_{f,m}(t)$, the backward reflection coefficient $\Gamma_{b,m}(t)$, and the regression coefficient $\kappa_{m,\lambda a}(t)$ register 90, 92, and 96 values.

Intermediate variables include a weighted sum of the forward prediction error squares $\Im_m(t)$, a weighted sum of the backward prediction error squares $\beta_m(t)$, a scalar parameter $\Delta_m(t)$, a conversion factor $\gamma_m(t)$, and another scalar parameter $\rho_{m,\lambda_a}(t)$. The weighted sum of the forward prediction errors $\Im_m(t)$ is defined as:

$$\Im_m(t) = \sum_{i=1}^{t} \lambda^{t-i} \left| f_m(i) \right|^2; \tag{44}$$

where λ without a wavelength identifier, a or b, is a constant multiplicative value unrelated to wavelength and is typically less than or equal to one, i.e., $\lambda \leq 1$. The weighted sum of the backward prediction errors $\beta_m(t)$ is defined as:

$$\beta_m(t) = \sum_{i=1}^{t} \lambda^{t-i} \left| b_m(i) \right|^2 \tag{45}$$

where, again, λ without a wavelength identifier, a or b, is a constant multiplicative value unrelated to wavelength and is typically less than or equal to one, i.e., $\lambda \leq 1$. These weighted sum intermediate error signals can be manipulated such that they are more easily solved for, as described in Chapter 9, § 9.3 of the Haykin book referenced above and defined hereinafter in equations (59) and (60).

DESCRIPTION OF THE JOINT PROCESS ESTIMATOR

The operation of the joint process estimator 60 is as follows. When the joint process estimator 60 is turned on, the initial values of intermediate variables and signals including the parameter $\Delta_{m-1}(t)$, the weighted sum of the forward prediction error signals $\Im_{m-1}(t)$, the weighted sum of the backward prediction error signals $\beta_{m-1}(t)$, the parameter

-48-

$\rho_{m,\lambda a}(t)$, and the zero-stage estimation error $e_{0,\lambda a}(t)$ are initialized, some to zero and some to a small positive number δ:

$$\Delta_{m-1}(0) = 0; \tag{46}$$

$$\vartheta_{m-1}(0) = \delta; \tag{47}$$

$$\beta_{m-1}(0) = \delta; \tag{48}$$

$$\rho_{m,\lambda a}(0) = 0; \tag{49}$$

$$e_{0,\lambda a}(t) = S_{\lambda a}(t) \text{ for } t \geq 0. \tag{50}$$

After initialization, a simultaneous sample of the measured signal $S_{\lambda a}(t)$ or $S_{\lambda b}(t)$ and either the secondary reference $n'(t)$ or the primary reference $s'(t)$ are input to the joint process estimator 60, as shown in Figure 8. The forward and backward prediction error signals $f_0(t)$ and $b_0(t)$, and intermediate variables including the weighted sums of the forward and backward error signals $\vartheta_0(t)$ and $\beta_0(t)$, and the conversion factor $\gamma_0(t)$ are calculated for the zero-stage according to:

$$f_0(t) = b_0(t) = n'(t) \tag{51a}$$

$$\vartheta_0(t) = \beta_0(t) = \lambda \vartheta_0(t-1) + |n'(t)|^2 \tag{52a}$$

$$\gamma_0(t-1) = 1 \tag{53a}$$

if a secondary reference $n'(t)$ is used or according to:

$$f_0(t) = b_0(t) = s'(t) \tag{51b}$$

$$\vartheta_0(t) = \beta_0(t) = \lambda \vartheta_0(t-1) + |s'(t)|^2 \tag{52b}$$

-49-

$$\Upsilon_0(t-1) = 1 \qquad\qquad (53b)$$

if a primary reference s'(t) is used where, again, λ without a wavelength identifier, a or b, is a constant multiplicative value unrelated to wavelength.

Forward reflection coefficient $\Gamma_{f,m}(t)$, backward reflection coefficient $\Gamma_{b,m}(t)$, and regression coefficient $\kappa_{m,\lambda a}(t)$ register 90, 92 and 96 values in each stage thereafter are set according to the output of the previous stage. The forward reflection coefficient $\Gamma_{f,1}(t)$, backward reflection coefficient $\Gamma_{b,1}(t)$, and regression coefficient $\kappa_{1,\lambda a}(t)$ register 90, 92 and 96 values in the first stage are thus set according to the algorithm using values in the zero-stage of the joint process estimator 60. In each stage, m≥1, intermediate values and register values including the parameter $\Delta_{m-1}(t)$; the forward reflection coefficient $\Gamma_{f,m}(t)$ register 90 value; the backward reflection coefficient $\Gamma_{b,m}(t)$ register 92 value; the forward and backward error signals $f_m(t)$ and $b_m(t)$; the weighted sum of squared forward prediction errors $\Im_{f,m}(t)$, as manipulated in § 9.3 of the Haykin book; the weighted sum of squared backward prediction errors $\beta_{b,m}(t)$, as manipulated in § 9.3 of the Haykin book; the conversion factor $\gamma_m(t)$; the parameter $\rho_{m,\lambda a}(t)$; the regression coefficient $\kappa_{m,\lambda a}(t)$ register 96 value; and the estimation error $e_{m+1\lambda a}(t)$ value are set according to:

$$\Delta_{m-1}(t) = \lambda \, \Delta_{m-1}(t-1) + \{b_{m-1}(t-1)f^*_{m-1}(t)/\gamma_{m-1}(t-1)\} \qquad (54)$$

$$\Gamma_{f,m}(t) = -\{\Delta_{m-1}(t)/\beta_{m-1}(t-1)\} \qquad (55)$$

$$\Gamma_{b,m}(t) = -\{\Delta^*_{m-1}(t)/\Im_{m-1}(t)\} \qquad (56)$$

$$f_m(t) = f_{m-1}(t) + \Gamma^*_{f,m}(t)b_{m-1}(t-1) \qquad (57)$$

$$b_m(t) = b_{m-1}(t-1) + \Gamma^*_{b,m}(t)f_{m-1}(t) \qquad (58)$$

-50-

$$\Im_m(t) = \Im_{m-1}(t) - \left\{ |\Delta_{m-1}(t)|^2 / \mathcal{B}_{m-1}(t-1) \right\} \qquad (59)$$

$$\mathcal{B}_m(t) = \mathcal{B}_{m-1}(t-1) - \left\{ |\Delta_{m-1}(t)|^2 / \Im_{m-1}(t) \right\} \qquad (60)$$

$$\gamma_m(t-1) = \gamma_{m-1}(t-1) - \left\{ |b_{m-1}(t-1)|^2 / \mathcal{B}_{m-1}(t-1) \right\} \qquad (61)$$

$$\rho_{m,\lambda a}(t) = \lambda \rho_{m,\lambda a}(t-1) + \left\{ b_m(t) \epsilon^*_{m,\lambda a}(t) / \gamma_m(t) \right\} \qquad (62)$$

$$\kappa_{m,\lambda a}(t) = \left\{ \rho_{m,\lambda a}(t) / \mathcal{B}_m(t) \right\} \qquad (63)$$

$$\epsilon_{m+1,\lambda a}(t) = \epsilon_{m,\lambda a}(t) - \kappa^*_m(t) b_m(t) \qquad (64)$$

where a (*) denotes a complex conjugate.

These equations cause the error signals $f_m(t)$, $b_m(t)$, $e_{m,\lambda a}(t)$ to be squared or to be multiplied by one another, in effect squaring the errors, and creating new intermediate error values, such as $\Delta_{m-1}(t)$. The error signals and the intermediate error values are recursively tied together, as shown in the above equations (54) through (64). They interact to minimize the error signals in the next stage.

After a good approximation to either the primary signal $s_{\lambda a}(t)$ or the secondary signal $n_{\lambda a}(t)$ has been determined by the joint process estimator 60, a next set of samples, including a sample of the measured signal $S_{\lambda a}(t)$ and a sample of either the secondary reference n'(t) or the primary reference s'(t), are input to the joint process estimator 60. The re-initialization process does not re-occur, such that the forward and backward reflection coefficient $\Gamma_{f,m}(t)$ and $\Gamma_{b,m}(t)$ register 90, 92 values and the regression coefficient $\kappa_{m,\lambda a}(t)$ register 96 value reflect the multiplicative values required to estimate either the primary signal portion $s_{\lambda a}(t)$ or the secondary signal portion $n_{\lambda a}(t)$ of the sample of $S_{\lambda a}(t)$ input previously. Thus, information from previous samples is used to estimate either the primary or secondary signal portion of a present set of samples in each stage.

-51-

In a more numerically stable and preferred embodiment of the above described joint process estimator, a normalized joint process estimator is used. This version of the joint process estimator normalizes several variables of the above-described joint process estimator such that the normalized variables fall between -1 and 1. The derivation of the normalized joint process estimator is motivated in the Haykin text as problem 12 on page 640 by redefining the variables defined according to the following conditions:

$$\overline{f}_m(t) = \frac{f_m(t)}{\sqrt{\mathscr{F}_m(t)\gamma_m(t-1)}}$$

$$\overline{b}(t) = \frac{b_m(t)}{\sqrt{\beta_m(t)\gamma_m(t)}}$$

$$\overline{\Delta}_m(t) = \frac{\Delta_m(t)}{\sqrt{\mathscr{F}_m(t)\beta_m(t-1)}}$$

This transformation allows the conversion of Equations (54) - (64) to the following normalized equations:

$$\overline{\Delta}_{m-1}(t) = \overline{\Delta}_{m-1}(t-1)\left[1-|f_{m-1}(t)|^2\right]^{1/2}\left[1-|\overline{b}_{m-1}(t-1)|^2\right]^{1/2} + \overline{b}_{m-1}(t-1)\overline{f}_{m-1}(t)$$

$$\overline{b}_m(t) = \frac{\left[\overline{b}_{m-1}(t-1) - \overline{\Delta}_{m-1}(t)\overline{f}_{m-1}(t)\right]}{\left[1-|\overline{\Delta}_{m-1}(t)|^2\right]^{1/2}\left[1-|\overline{f}_{m-1}(t)|^2\right]^{1/2}} \cdots$$

$$\overline{f}_m(t) = \frac{\left[\overline{f}_{m-1}(t) - \overline{\Delta}_{m-1}(t)\overline{b}_{m-1}(t-1)\right]}{\left[1-|\overline{\Delta}_{m-1}(t)|^2\right]^{1/2}\left[1-|\overline{b}_{m-1}(t-1)|^2\right]^{1/2}}$$

$$\beta_m(t) = \left[1-|\overline{\Delta}_{m-1}(t)|^2\right]\beta_{m-1}(t-1)$$

$$\gamma_m(t) = \gamma_{m-1}(t)\left[1-|\overline{b}_{m-1}(t)|^2\right]$$

$$\rho_m(t) = \lambda \cdot \left[\frac{\gamma_m(t)\beta_m(t-1)}{\gamma_m(t-1)\beta_m(t)} \right]^{1/2} \rho_m(t-1) + \overline{b}_m(t)\,\varepsilon_m(t)$$

$$\varepsilon_{m+1,\lambda_a}(t) = \varepsilon_{m,\lambda_a}(t) - \rho_m(t)\,\overline{b}_m(t)$$

Initialization of Normalized Joint Process Estimator

Let $N(t)$ be defined as the reference noise input at time index n and $U(t)$ be defined as combined signal plus noise input at time index t the following equations apply (see Haykin, p. 619):

1.   To initialize the algorithm, at time t=0 set

$$\overline{\Delta}_{m-1}(0) = 0$$

$$\beta_{m-1}(0) = \delta = 10^{-6}$$

$$\gamma_0(0) = 1$$

2.   At each instant t≥1, generate the various zeroth-order variables as follows:

$$\gamma_0(t-1) = 1$$

$$\beta_0(t) = \lambda\beta_0(t-1) + N^2(t)$$

$$\overline{b}_0(t) = \overline{f}_0(t) = \frac{N(t)}{\sqrt{\beta_0(t)}}$$

3.   For regression filtering, initialize the algorithm by setting at time index t=0

-53-

$$\rho_m(0) = 0$$

4. At each instant t≥1, generate the zeroth-order variable

$$\varepsilon_0(t) = U(t).$$

Accordingly, a normalized joint process estimator can be used for a more stable system.

In yet another embodiment, the correlation cancellation is performed with a QRD algorithm as shown diagrammatically in Figure 8a and as described in detail in Chapter 18 of Adaptive Filter Theory by Simon Haykin, published by Prentice-Hall, ~~copyright 1986~~.

The following equations adapted from the Haykin book correspond ~~to the QRD-LSL diagram of Figure 8a (also adapted~~ from the Haykin book).

Computations

a. Predictions: For time t = 1, 2, ..., and prediction order m = 1, 2, ..., M, where M is the final prediction order, compute:

$$\beta_{m-1}(t-1) = \lambda\beta_{m-1}(t-2) + |\epsilon_{b,m-1}(t-1)|^2$$

$$c_{b,m-1}(t+1) = \frac{\lambda^{1/2}\beta_{m-1}^{1/2}(t-2)}{\beta_{m-1}^{1/2}(t-1)}$$

$$s_{b,m-1}(t-1) = \frac{\dot{\epsilon}_{b,m-1}(t-1)}{\beta_{m-1}^{1/2}(t-1)}$$

$$\epsilon_{f,m}(t) = c_{b,m-1}(t-1)\epsilon_{f,m-1}(t) - \dot{s}_{b,m-1}(t-1)\lambda^{1/2}\dot{\pi}_{f,m-1}(t-1)$$

$$\dot{\pi}_{f,m-1}(t) = c_{b,m-1}(t-1)\lambda^{1/2}\dot{\pi}_{f,m-1}(t-1) + s_{b,m-1}(t-1)\epsilon_{f,m-1}(t)$$

$$\gamma_m^{1/2}(t-1) = c_{b,m-1}(t-1)\gamma_{m-1}^{1/2}(t-1)$$

$$\mathscr{F}_{m-1}(t) = \lambda\mathscr{F}_{m-1}(t-1) + |\epsilon_{f,m-1}(t)|^2$$

$$c_{f,m-1}(t) = \frac{\lambda^{1/2}\mathscr{F}_{m-1}^{1/2}(t-1)}{\mathscr{F}_{m-1}^{1/2}(t)}$$

$$s_{f,m-1}(t) = \frac{\dot{\epsilon}_{f,m-1}(t)}{\mathscr{F}_{m-1}^{1/2}(t)}$$

$$\epsilon_{b,m}(t) = c_{f,m-1}(t)\epsilon_{b,m-1}(t-1) - \dot{s}_{f,m-1}(t)\lambda^{1/2}\dot{\pi}_{b,m-1}(t-1)$$

$$\dot{\pi}_{b,m-1}(t) = c_{f,m-1}(t)\lambda^{1/2}\dot{\pi}_{b,m-1}(t-1) + s_{f,m-1}(t)\epsilon_{b,m-1}(t-1)$$

b.   Filtering: For order m = 0, 1, ... , M-1; and time t=1, 2, ..., compute

$$\beta_m(t) = \lambda\beta_m(t-1) + |\epsilon_{b,m}(t)|^2$$

5

$$c_{b,m}(t) = \frac{\lambda^{1/2}\beta_m^{1/2}(t-1)}{\beta_m^{1/2}(t)}$$

$$s_{b,m}(t) = \frac{\epsilon_{b,m}^{\cdot}(t)}{\beta_m^{1/2}(t)}$$

$$\epsilon_{m+1}(t) = c_{b,m}(t)\epsilon_m(t) - s_{b,m}^{\cdot}(t)\lambda^{1/2}\rho_m^{\cdot}(t-1)$$

$$\rho_m^{\cdot}(t) = c_{b,m}(t)\lambda^{1/2}\rho_m^{\cdot}(t-1) + s_{b,m}(t)\epsilon_m(t)$$

$$\gamma_{m+1}^{1/2}(t) = c_{b,m}(t)\gamma_m^{1/2}(t)$$

$$\mathcal{E}_{m+1}(t) = \lambda_{m+1}^{1/2}(t)\epsilon_{m+1}(t)$$

5 5. <u>Initialization</u>

　　a. <u>Auxiliary parameter initialization</u>: for order m = 1, 2, ..., M, set

$$\pi_{f,m-1}(0) = \pi_{b,m-1}(0) = 0$$

$$p_m(0) = 0$$

　　b. <u>Soft constraint initialization</u>: For order m = 0,
10 1, ..., M, set

$$\beta_m(-1) = \delta$$

$$\mathcal{F}_m(0) = \delta$$

where δ is a small positive constant.

　　c. <u>Data initialization</u>: For t = 1, 2, ..., compute

$$\epsilon_{f,0}(t) = \epsilon_{b,0}(t) = \mu(t)$$

$$\epsilon_0(t) = d(t)$$

$$\gamma_0(t) = 1$$

where $\mu(t)$ is the input and d(t) is the desired response at time t.

FLOWCHART OF JOINT PROCESS ESTIMATOR

5   In a signal processor, such as a physiological monitor incorporating a reference processor of the present invention to determine a reference n'(t) or s'(t) for input to a correlation canceler, a joint process estimator 60 type adaptive noise canceler is generally implemented via a
10   software program having an iterative loop. One iteration of the loop is analogous to a single stage of the joint process estimator as shown in Figure 8. Thus, if a loop is iterated m times, it is equivalent to an m stage joint process estimator 60.

15   A flow chart of a subroutine to estimate the primary signal portion $s_{\lambda a}(t)$ or the secondary signal portion $n_{\lambda a}(t)$ of a measured composite signal, $S_{\lambda a}(t)$ is shown in Figure 9. The flow chart illustrates the function of a reference processor for determining either the secondary reference
20   n'(t) or the primary reference s'(t). The flowchart for the joint process estimator is implemented in software.

A one-time initialization is performed when the physiological monitor is powered-on, as indicated by an "INITIALIZE NOISE CANCELER" action block 120. The
25   initialization sets all registers 90, 92, and 96 and delay element variables 110 to the values described above in equations (46) through (50).

Next, a set of simultaneous samples of the composite measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ is input to the subroutine
30   represented by the flowchart in Figure 9. Then a time update

-57-

of each of the delay element program variables occurs, as indicated in a "TIME UPDATE OF [Z^{-1}] ELEMENTS" action block 130. The value stored in each of the delay element variables 110 is set to the value at the input of the delay element variable 110. Thus, the zero-stage backward prediction error $b_0(t)$ is stored as the first-stage delay element variable, and the first-stage backward prediction error $b_1(t)$ is stored as the second-stage delay element variable, and so on.

Then, using the set of measured signal samples $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$, the reference signal is obtained using the ratiometric or the constant saturation methods described above. This is indicated by a "CALCULATE REFERENCE [n'(t) or s'(t)] FOR TWO MEASURED SIGNAL SAMPLES" action block 140.

A zero-stage order update is performed next as indicated in a "ZERO-STAGE UPDATE" action block 150. The zero-stage backward prediction error $b_0(t)$, and the zero-stage forward prediction error $f_0(t)$ are set equal to the value of the reference signal n'(t) or s'(t). Additionally, the weighted sum of the forward prediction errors $\Im_m(t)$ and the weighted sum of backward prediction errors $\beta_m(t)$ are set equal to the value defined in equations (47) and (48).

Next, a loop counter, m, is initialized as indicated in a "m=0" action block 160. A maximum value of m, defining the total number of stages to be used by the subroutine corresponding to the flowchart in Figure 9, is also defined. Typically, the loop is constructed such that it stops iterating once a criterion for convergence upon a best approximation to either the primary signal or the secondary signal has been met by the joint process estimator 60. Additionally, a maximum number of loop iterations may be chosen at which the loop stops iteration. In a preferred embodiment of a physiological monitor of the present invention, a maximum number of iterations, m=6 to m=10, is advantageously chosen.

Within the loop, the forward and backward reflection coefficient $\Gamma_{f,m}(t)$ and $\Gamma_{b,m}(t)$ register 90 and 92 values in

the least-squares lattice filter are calculated first, as indicated by the "ORDER UPDATE MTH CELL OF LSL-LATTICE" action block 170 in Figure 9. This requires calculation of intermediate variable and signal values used in determining

5   register 90, 92, and 96 values in the present stage, the next stage, and in the regression filter 80.

The calculation of regression filter register 96 value $\kappa_{m,\lambda a}(t)$ is performed next, indicated by the "ORDER UPDATE MTH STAGE OF REGRESSION FILTER(S)" action block 180. The two

10   order update action blocks 170 and 180 are performed in sequence m times, until m has reached its predetermined maximum (in the preferred embodiment, m=6 to m=10) or a solution has been converged upon, as indicated by a YES path from a "DONE" decision block 190. In a computer subroutine,

15   convergence is determined by checking if the weighted sums of the forward and backward prediction errors $\mathfrak{F}_m(t)$ and $\mathfrak{B}_m(t)$ are less than a small positive number. An output is calculated next, as indicated by a "CALCULATE OUTPUT" action block 200. The output is a good approximation to either the primary

20   signal or secondary signal, as determined by the reference processor 26 and joint process estimator 60 subroutine corresponding to the flow chart of Figure 9. This is displayed (or used in a calculation in another subroutine), as indicated by a "TO DISPLAY" action block 210.

25   A new set of samples of the two measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ is input to the processor and joint process estimator 60 adaptive noise canceler subroutine corresponding to the flowchart of Figure 9 and the process reiterates for these samples. Note, however, that the initialization

30   process does not re-occur. New sets of measured signal samples $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ are continuously input to the reference processor 26 and joint process estimator adaptive noise canceler subroutine. The output forms a chain of samples which is representative of a continuous wave. This

35   waveform is a good approximation to either the primary signal waveform $s_{\lambda a}(t)$ or the secondary waveform $n_{\lambda a}(t)$ at

-59-

wavelength λa. The waveform may also be a good approximation to either the primary signal waveform $s_{\lambda b}(t)$ or the secondary waveform $n"_{\lambda b}(t)$ at wavelength λb.

A corresponding flowchart for the QRD algorithm of Figure 8a is depicted in Figure 9a, with reference numeral corresponding in number with an 'a' extension

CALCULATION OF SATURATION FROM
CORRELATION CANCELER OUTPUT

Physiological monitors may use the approximation of the primary signals $s"_{\lambda a}(t)$ or $s"_{\lambda b}(t)$ or the secondary signals $n"_{\lambda a}(t)$ or $n"_{\lambda b}(t)$ to calculate another quantity, such as the saturation of one constituent in a volume containing that constituent plus one or more other constituents. Generally, such calculations require information about either a primary or secondary signal at two wavelengths. For example, the constant saturation method requires a good approximation of the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ of both measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$. The arterial saturation is determined from the approximations to both signals, i.e. $s"_{\lambda a}(t)$ and $s"_{\lambda b}(t)$. The constant saturation method also requires a good approximation of the secondary signal portions $n_{\lambda a}(t)$ or $n_{\lambda b}(t)$. An estimate of the venous saturation may be determined from the approximations to these signals i. e. $n"_{\lambda a}(t)$ and $n"_{\lambda b}(t)$.

A joint process estimator 60 having two regression filters 80a and 80b is shown in Figure 10. A first regression filter 80a accepts a measured signal $S_{\lambda a}(t)$. A second regression filter 80b accepts a measured signal $S_{\lambda b}(t)$ for a use of the constant saturation method to determine the reference signal $n'(t)$ or $s'(t)$. The first and second regression filters 80a and 80b are independent. The backward prediction error $b_m(t)$ is input to each regression filter 80a and 80b, the input for the second regression filter 80b bypassing the first regression filter 80a.

The second regression filter 80b comprises registers 98, and summing elements 108 arranged similarly to those in the

first regression filter 80a. The second regression filter 80b operates via an additional intermediate variable in conjunction with those defined by equations (54) through (64), i.e.:

5

$$\rho_{m,\lambda b}(t) = \lambda \rho_{m,\lambda b}(t-1) + \{b_m(t)e^*_{m,\lambda b}(t)/\gamma_m(t)\}; \quad \text{and} \quad (65)$$

$$\rho_{0,\lambda b}(0) = 0. \quad (66)$$

10 The second regression filter 80b has an error signal value defined similar to the first regression filter error signal values, $e_{m+1,\lambda a}(t)$, i.e.:

$$e_{m+1,\lambda b}(t) = e_{m,\lambda b}(t) - \kappa^-_{m,\lambda b}(t)b_m(t); \quad \text{and} \quad (67)$$

15

$$e_{0,\lambda b}(t) = S_{\lambda b}(t) \text{ for } t \geq 0. \quad (68)$$

The second regression filter has a regression coefficient $\kappa_{m,\lambda b}(t)$ register 98 value defined similarly to the first 20 regression filter error signal values, i.e.:

$$\kappa_{m,\lambda b}(t) = \{\rho_{m,\lambda b}(t)/\beta_m(t)\}; \quad \text{or} \quad (69)$$

These values are used in conjunction with those intermediate 25 variable values, signal values, register and register values defined in equations (46) through (64). These signals are calculated in an order defined by placing the additional signals immediately adjacent a similar signal for the wavelength λa.

30 For the constant saturation method, $S_{\lambda b}(t)$ is input to the second regression filter 80b. The output is then a good approximation to the primary signal $s''_{\lambda b}(t)$ or secondary signal $s''_{\lambda b}(t)$.

The addition of the second regression filter 80b does 35 not substantially change the computer program subroutine represented by the flowchart of Figure 9. Instead of an

-61-

order update of the mth stage of only one regression filter, an order update of the mth stage of both regression filters 80a and 80b is performed. This is characterized by the plural designation in the "ORDER UPDATE OF mth STAGE OF REGRESSION FILTER(S)" activity block 180 in Figure 9. Since the regression filters 80a and 80b operate independently, independent calculations can be performed in the reference processor and joint process estimator 60 adaptive noise canceler subroutine modeled by the flowchart of Figure 9.

An alternative diagram for the joint process estimator of Figure 10, using the QRD algorithm and having two regression filters is shown in Figure 10a. This type of joint process estimator would be used for correlation cancellation using the QRD algorithm described in the Haykin book.

<u>CALCULATION OF SATURATION</u>

Once good approximations to the primary signal portions $s"_{\lambda a}(t)$ and $s"_{\lambda b}(t)$ or the secondary signal portion, $n"_{\lambda a}(t)$ and $n"_{\lambda b}(t)$, have been determined by the joint process estimator 60, the saturation of A_5 in a volume containing A_5 and A_6, for example, may be calculated according to various known methods. Mathematically, the approximations to the primary signals can be written in terms of λa and λb, as:

$$s"_{\lambda a}(t) \sim \epsilon_{5,\lambda a}c_5 x_{5,6}(t) + \epsilon_{6,\lambda a}c_6 x_{5,6}(t); \text{ and} \qquad (70)$$

$$s"_{\lambda b}(t) \sim \epsilon_{5,\lambda b}c_5 x_{5,6}(t) + \epsilon_{6,\lambda b}c_6 x_{5,6}(t). \qquad (71)$$

Equations (70) and (71) are equivalent to two equations having three unknowns, namely $c_5(t)$, $c_6(t)$ and $x_{5,6}(t)$. The saturation can be determined by acquiring approximations to the primary or secondary signal portions at two different, yet proximate times t_1 and t_2 over which the saturation of A_5 in the volume containing A_5 and A_6 and the saturation of A_3 in the volume containing A_3 and A_4 does not change

substantially.   For example, for the primary signals
estimated at times t_1 and t_2:

$$s''_{\lambda a}(t_1) \approx \epsilon_{s,\lambda a}c_5 x_{5,6}(t_1) + \epsilon_{6,\lambda a}c_6 x_{5,6}(t_1) \tag{72}$$

$$s''_{\lambda b}(t_1) \approx \epsilon_{s,\lambda b}c_5 x_{5,6}(t_1) + \epsilon_{6,\lambda b}c_6 x_{5,6}(t_1) \tag{73}$$

$$s''_{\lambda a}(t_2) \approx \epsilon_{s,\lambda a}c_5 x_{5,6}(t_2) + \epsilon_{6,\lambda a}c_6 x_{5,6}(t_2) \tag{74}$$

$$s''_{\lambda b}(t_2) \approx \epsilon_{s,\lambda b}c_5 x_{5,6}(t_2) + \epsilon_{6,\lambda b}c_6 x_{5,6}(t_2) \tag{75}$$

Then, difference signals may be determined which relate the
signals of equations (72) through (75), i.e.:

$$\Delta s_{\lambda a} = s''_{\lambda a}(t_1) - s''_{\lambda a}(t_2) \approx \epsilon_{s,\lambda a}c_5 \Delta x + \epsilon_{6,\lambda a}c_6 \Delta x; \text{ and} \tag{76}$$

$$\Delta s_{\lambda b} = s''_{\lambda b}(t_1) - s''_{\lambda b}(t_2) \approx \epsilon_{s,\lambda b}c_5 \Delta x + \epsilon_{,\lambda b}c_6 \Delta x; \tag{77}$$

where $\Delta x = x_{5,6}(t_1) - x_{5,6}(t_2)$. The average saturation at time
$t = (t_1 + t_2)/2$ is:

$$\text{Saturation}(t) = c_5(t)/[c_5(t) + c_6(t)] \tag{78}$$

$$= \frac{\epsilon_{6,\lambda a} - \epsilon_{6,\lambda b}(\Delta s_{\lambda a}/\Delta s_{\lambda b})}{\epsilon_{6,\lambda a} - \epsilon_{s,\lambda a} - (\epsilon_{6,\lambda b} - \epsilon_{s,\lambda b})(\Delta s_{\lambda a}/\Delta s_{\lambda b})} \tag{79}$$

It will be understood that the Δx term drops out from the
saturation calculation because of the division.   Thus,
knowledge of the thickness of the primary constituents is not
required to calculate saturation.

PULSE OXIMETRY MEASUREMENTS

A specific example of a physiological monitor utilizing
a processor of the present invention to determine a secondary
reference n'(t) for input to a correlation canceler that
removes erratic motion-induced secondary signal portions is

a pulse oximeter. Pulse oximetry may also be performed utilizing a processor of the present invention to determine a primary signal reference s'(t) which may be used for display purposes or for input to a correlation canceler to derive information about patient movement and venous blood oxygen saturation.

A pulse oximeter typically causes energy to propagate through a medium where blood flows close to the surface for example, an ear lobe, or a digit such as a finger, a forehead or a fetus' scalp. An attenuated signal is measured after propagation through or reflected from the medium. The pulse oximeter estimates the saturation of oxygenated blood.

Freshly oxygenated blood is pumped at high pressure from the heart into the arteries for use by the body. The volume of blood in the arteries varies with the heartbeat, giving rise to a variation in absorption of energy at the rate of the heartbeat, or the pulse.

Oxygen depleted, or deoxygenated, blood is returned to the heart by the veins along with unused oxygenated blood. The volume of blood in the veins varies with the rate of breathing, which is typically much slower than the heartbeat. Thus, when there is no motion induced variation in the thickness of the veins, venous blood causes a low frequency variation in absorption of energy. When there is motion induced variation in the thickness of the veins, the low frequency variation in absorption is coupled with the erratic variation in absorption due to motion artifact.

In absorption measurements using the transmission of energy through a medium, two light emitting diodes (LED's) are positioned on one side of a portion of the body where blood flows close to the surface, such as a finger, and a photodetector is positioned on the opposite side of the finger. Typically, in pulse oximetry measurements, one LED emits a visible wavelength, preferably red, and the other LED emits an infrared wavelength. However, one skilled in the art will realize that other wavelength combinations could be

used.   The finger comprises skin, tissue, muscle, both arterial blood and venous blood, fat, etc., each of which absorbs light energy differently due to different absorption coefficients,   different   concentrations,   different thicknesses, and changing optical pathlengths.   When the patient is not moving, absorption is substantially constant except for the flow of blood.   The constant attenuation can be determined and subtracted from the signal via traditional filtering techniques.   When the patient moves, this causes perturbation such as changing optical pathlength due to movement of background fluids (e.g., venous blood having a different saturation than the arterial blood).   Therefore, the measured signal becomes erratic.   Erratic motion induced noise typically cannot be predetermined and/or subtracted from   the   measured   signal   via   traditional   filtering techniques.   Thus, determining the oxygen saturation of arterial blood and venous blood becomes more difficult.

A   schematic   of   a   physiological   monitor   for   pulse oximetry   is   shown   in   Figures   11-13.   Figure   11   depicts   a general hardware block diagram of a pulse oximeter 299.. A sensor 300 has two light emitters 301 and 302 such as LED's. One LED 301 emitting light of red wavelengths and another LED 302   emitting   light   of   infrared   wavelengths   are   placed adjacent   a   finger   310.   A   photodetector   320,   which   produces an electrical signal corresponding to the attenuated visible and   infrared   light   energy   signals   is   located   opposite   the LED's 301 and 302.   The photodetector 320 is connected to front end analog signal conditioning circuity 330.

The front end analog signal conditioning circuitry 330 has outputs coupled to analog to digital conversion circuit 332.   The   analog   to   digital   conversion   circuitry   332   has outputs   coupled   to   a   digital   signal   processing   system   334. The digital signal processing system 334 provides the desired parameters as outputs for a display 336.   Outputs for display are,   for   example,   blood   oxygen   saturation,   heart   rate,   and   a clean plethysmographic waveform.

The signal processing system also provides an emitter current control output 337 to a digital-to-analog converter circuit 338 which provides control information for light emitter drivers 340. The light emitter drivers 340 couple to

5   the light emitters 301, 302. The digital signal processing system 334 also provides a gain control output 342 for the front end analog signal conditioning circuitry 330.

Figure 11a illustrates a preferred embodiment for the combination of the emitter drivers 340 and the digital to

10   analog conversion circuit 338. As depicted in Figure 11a, the driver comprises first and second input latches 321, 322, a synchronizing latch 323, a voltage reference 324, a digital to analog conversion circuit 325, first and second switch banks 326, 327, first and second voltage to current

15   converters 328, 329 and the LED emitters 301, 302 corresponding to the LED emitters 301, 302 of Figure 11.

The preferred driver depicted in Figure 11a is advantageous in that the present inventors recognized that much of the noise in the oximeter 299 of Figure 11 is caused

20   by the LED emitters 301, 302. Therefore, the emitter driver circuit of Figure 11a is designed to minimize the noise from the emitters 301, 302. The first and second input latches 321, 324 are connected directly to the DSP bus. Therefore, these latches significantly minimizes the bandwidth

25   (resulting in noise) present on the DSP bus which passes through to the driver circuitry of Figure 11a. The output of the first and second input latches only changes when these latched detect their address on the DSP bus. The first input latch receives the setting for the digital to analog

30   converter circuit 325. The second input latch receives switching control data for the switch banks 326, 327. The synchronizing latch accepts the synchronizing pulses which maintain synchronization between the activation of emitters 301, 302 and the analog to digital conversion circuit 332.

35   The voltage reference is also chosen as a low noise DC voltage reference for the digital to analog conversion

-66-

circuit 325. In addition, in the present embodiment, the
voltage reference has an lowpass output filter with a very
low corner frequency (e.g., 1Hz in the present embodiment).
The digital to analog converter 325 also has a lowpass filter
5   at its output with a very low corner frequency (e.g., 1 Hz).
The digital to analog converter provides signals for each of
the emitters 301, 302.

In the present embodiment, the output of the voltage to
current converters 328, 329 are switched such that with the
10   emitters 301, 302 connected in back-to-back configuration,
only one emitter is active an any given time. In addition,
the voltage to current converter for the inactive emitter is
switched off at its input as well, such that it is completely
deactivated.   This reduces noise from the switching and
15   voltage to current conversion circuitry. In the present
embodiment, low noise voltage to current converters are
selected (e.g., Op 27 Op Amps), and the feedback loop is
configured to have a low pass filter to reduce noise. In the
present embodiment, the low pass filtering function of the
20   voltage to current converter 328, 329 has a corner frequency
of just above 625 Hz, which is the switching speed for the
emitters, as further discussed below. Accordingly, the
preferred driver circuit of Figure 11a, minimizes the noise
of the emitters 301, 302.

25   In general, the red and infrared light emitters 301, 302
each emits energy which is absorbed by the finger 310 and
received by the photodetector 320. The photodetector 320
produces an electrical signal which corresponds to the
intensity of the light energy striking the photodetector 320.
30   The front end analog signal conditioning circuitry 330
receives the intensity signals and filters and conditions
these signals as further described below for further
processing. The resultant signals are provided to the
analog-to-digital conversion circuitry 332 which converts the
35   analog signals to digital signals for further processing by
the digital signal processing system 334. The digital signal

processing system 334 utilizes the two signals in order to provide a what will be called herein a "saturation transform." It should be understood, that for parameters other than blood saturation monitoring, the saturation transform could be better termed as a concentration transform, in-vivo transform, or the like, depending on the desired parameter. The term saturation transform is used to describe an operation which converts the sample data from time domain to saturation domain values as will be apparent from the discussion below. In the present embodiment, the output of the digital signal processing system 334 provides clean plethysmographic waveforms of the detected signals and provides values for oxygen saturation and pulse rate to the display 336.

It should be understood that in different embodiments of the present invention, one or more of the outputs may be provided. The digital signal processing system 334 also provides control for driving the light emitters 301, 302 with an emitter current control signal on the emitter current control output 337. This value is a digital value which is converted by the digital-to-analog conversion circuit 338 which provides a control signal to the emitter current drivers 340. The emitter current drivers 340 provide the appropriate current drive for the red emitter 301 and the infrared emitter 302. Further detail of the operation of the physiological monitor for pulse oximetry is explained below.

In the present embodiment, the light emitters are driven via the emitter current driver 340 to provide light transmission with digital modulation at 625 Hz. In the present embodiment, the light emitters 301, 302 are driven at a power level which provides an acceptable intensity for detection by the detector and for conditioning by the front end analog signal conditioning circuitry 330. Once this energy level is determined for a given patient by the digital signal processing system 334, the current level for the red and infrared emitters is maintained constant. It should be

understood, however, that the current could be adjusted for changes in the ambient room light and other changes which would effect the voltage input to the front end analog signal conditioning circuitry 330. In the present invention, the red and infrared light emitters are modulated as follows: for one complete 625 Hz red cycle, the red emitter 301 is activated for the first quarter cycle, and off for the remaining three-quarters cycle; for one complete 625 Hz infrared cycle, the infrared light emitter 302 is activated for one quarter cycle and is off for the remaining three-quarters cycle. In order to only receive one signal at a time, the emitters are cycled on and off alternatively, in sequence, with each only active for a quarter cycle per 625 Hz cycle and a quarter cycle separating the active times.

The light signal is attenuated (amplitude modulated) by the pumping of blood through the finger 310 (or other sample medium). The attenuated (amplitude modulated) signal is detected by the photodetector 320 at the 625 Hz carrier frequency for the red and infrared light. Because only a single photodetector is used, the photodetector 320 receives both the red and infrared signals to form a composite time division signal.

The composite time division signal is provided to the front analog signal conditioning circuitry 330. Additional detail regarding the front end analog signal conditioning circuitry 330 and the analog to digital converter circuit 332 is illustrated in Figure 12. As depicted in Figure 12, the front end circuity 302 has a preamplifier 342, a high pass filter 344, an amplifier 346, a programmable gain amplifier 348, and a low pass filter 350. The preamplifier 342 is a transimpedance amplifier that converts the composite current signal from the photodetector 320 to a corresponding voltage signal, and amplifies the signal. In the present embodiment, the preamplifier has a predetermined gain to boost the signal amplitude for ease of processing. In the present embodiment, the source voltages for the preamplifier 342 are -15 VDC and

-69-

+15 VDC. As will be understood, the attenuated signal contains a component representing ambient light as well as the component representing the infrared or the red light as the case may be in time. If there is light in the vicinity of the sensor 300 other than the red and infrared light, this ambient light is detected by the photodetector 320. Accordingly, the gain of the preamplifier is selected in order to prevent the ambient light in the signal from saturating the preamplifier under normal and reasonable operating conditions.

In the present embodiment, the preamplifier 342 comprises an Analog Devices AD743JR OpAmp. This transimpedance amplifier is particularly advantageous in that it exhibits several desired features for the system described, such as: low equivalent input voltage noise, low equivalent input current noise, low input bias current, high gain bandwidth product, low total harmonic distortion, high common mode rejection, high open loop gain, and a high power supply rejection ratio.

The output of the preamplifier 342 couples as an input to the high pass filter 344. The output of the preamplifier also provides a first input 346 to the analog to digital conversion circuit 332. In the present embodiment, the high pass filter is a single-pole filter with a corner frequency of about 1/2 - 1 Hz. However, the corner frequency is readily raised to about 90 Hz in one embodiment. As will be understood, the 625 Hz carrier frequency of the red and infrared signals is well above a 90 Hz corner frequency. The high-pass filter 344 has an output coupled as an input to an amplifier 346. In the present embodiment, the amplifier 346 comprises a unity gain amplifier. However, the gain of the amplifier 346 is adjustable by the variation of a single resistor. The gain of the amplifier 346 would be increased if the gain of the preamplifier 342 is decreased to compensate for the effects of ambient light.

-70-

The output of the amplifier 346 provides an input to a programmable gain amplifier 348. The programmable gain amplifier 348 also accepts a programming input from the digital signal processing system 334 on a gain control signal line 343. The gain of the programmable gain amplifier 348 is digitally programmable. The gain is adjusted dynamically at initialization or sensor placement for changes in the medium under test from patient to patient. For example, the signal from different fingers differs somewhat. Therefore, a dynamically adjustable amplifier is provided by the programmable gain amplifier 348 in order to obtain a signal suitable for processing.

The programmable gain amplifier is also advantageous in an alternative embodiment in which the emitter drive current is held constant. In the present embodiment, the emitter drive current is adjusted for each patient in order to obtain the proper dynamic range at the input of the analog to digital conversion circuit 332. However, changing the emitter drive current can alter the emitter wavelength, which in turn affects the end result in oximetry calculations. Accordingly, it would be advantageous to fix the emitter drive current for all patients. In an alternative embodiment of the present invention, the programmable gain amplifier can be adjusted by the DSP in order to obtain a signal at the input to the analog to digital conversion circuit which is properly within the dynamic range (+ 3v to -3v in the present embodiment) of the analog to digital conversion circuit 332. In this manner, the emitter drive current could be fixed for all patients, eliminating the wavelength shift due to emitter current drive changes.

The output of the programmable gain amplifier 348 couples as an input to a low-pass filter 350. Advantageously, the low pass filter 350 is a single-pole filter with a corner frequency of approximately 10 Khz in the present embodiment. This low pass filter provides anti-aliasing in the present embodiment.

The output of the low-pass filter 350 provides a second input 352 to the analog-to-digital conversion circuit 332. Figure 12 also depicts additional defect of the analog-to-digital conversion circuit. In the present embodiment, the

5 analog-to-digital conversion circuit 332 comprises a first analog-to-digital converter 354 and a second analog-to-digital converter 356. Advantageously, the first analog-to-digital converter 354 accepts input from the first input 346 to the analog-to-digital conversion circuit 332, and the

10 second analog to digital converter 356 accepts input on the second input 352 to the analog-to-digital conversion circuitry 332.

In one advantageous embodiment, the first analog-to-digital converter 354 is a diagnostic analog-to-digital

15 converter. The diagnostic task (performed by the digital signal processing system) is to read the output of the detector as amplified by the preamplifier 342 in order to determine if the signal is saturating the input to the high-pass filter 344. In the present embodiment, if the input to

20 the high pass filter 344 becomes saturated, the front end analog signal conditioning circuits 330 provides a '0' output. Alternatively, the first analog-to-digital converter 354 remains unused.

The second analog-to-digital converter 352 accepts the

25 conditioned composite analog signal from the front end signal conditioning circuitry 330 and converts the signal to digital form. In the present embodiment, the second analog to digital converter 356 comprises a single-channel, delta-sigma converter. In the present embodiment, a Crystal

30 Semiconductor CS5317-KS delta-sigma analog to digital converter is used. Such a converter is advantageous in that it is low cost, and exhibits low noise characteristics. More specifically, a delta-sigma converter consists of two major portions, a noise modulator and a decimation filter. The

35 selected converter uses a second order analog delta-sigma modulator to provide noise shaping. Noise shaping refers to

-72-

changing the noise spectrum from a flat response to a response where noise at the lower frequencies has been reduced by increasing noise at higher frequencies. The decimation filter then cuts out the reshaped, higher frequency noise to provide 16-bit performance at a lower frequency. The present converter samples the data 128 times for every 16 bit data word that it produces. In this manner, the converter provides excellent noise rejection, dynamic range and low harmonic distortion, that help in critical measurement situations like low perfusion and electrocautery.

In addition, by using a single-channel converter, there is no need to tune two or more channels to each other. The delta-sigma converter is also advantageous in that it exhibits noise shaping, for improved noise control. An exemplary analog to digital converter is a Crystal Semiconductor CS5317. In the present embodiment, the second analog to digital converter 356 samples the signal at a 20 Khz sample rate. The output of the second analog to digital converter 356 provides data samples at 20 Khz to the digital signal processing system 334 (Figure 11).

The digital signal processing system 334 is illustrated in additional detail in Figure 13. In the present embodiment, the digital signal processing system comprises a microcontroller 360, a digital signal processor 362, a program memory 364, a sample buffer 366, a data memory 368, a read only memory 370 and communication registers 372. In the present embodiment, the digital signal processor 362 is an Analog Devices AD 21020. In the present embodiment, the microcontroller 360 comprises a Motorola 68HC05, with built in program memory. In the present embodiment, the sample buffer 366 is a buffer which accepts the 20 Khz sample data from the analog to digital conversion circuit 332 for storage in the data memory 368. In the present embodiment, the data memory 368 comprises 32 KWords (words being 40 bits in the present embodiment) of static random access memory.

-73-

The microcontroller 360 is connected to the DSP 362 via a conventional JTAG Tap line. The microcontroller 360 transmits the boot loader for the DSP 362 to the program memory 364 via the Tap line, and then allows the DSP 362 to boot from the program memory 364. The boot loader in program memory 364 then causes the transfer of the operating instructions for the DSP 362 from the read only memory 370 to the program memory 364. Advantageously, the program memory 364 is a very high speed memory for the DSP 362.

The microcontroller 360 provides the emitter current control and gain control signals via the communications register 372.

Figures 14-20 depict functional block diagrams of the operations of the pulse oximeter 299 carried out by the digital signal processing system 334. The signal processing functions described below are carried out by the DSP 362 in the present embodiment with the microcontroller 360 providing system management. In the present embodiment, the operation is software/firmware controlled. Figure 14 depicts a generalized functional block diagram for the operations performed on the 20 Khz sample data entering the digital signal processing system 334. As illustrated in Figure 14, a demodulation, as represented in a demodulation module 400, is first performed. Decimation, as represented in a decimation module 402 is then performed on the resulting data. Certain statistics are calculated, as represented in a statistics module 404 and a saturation transform is performed, as represented in a saturation transform module 406, on the data resulting from the decimation operation. The data subjected to the statistics operations and the data subjected to the saturation transform operations are forwarded to saturation operations, as represented by a saturation calculation module 408 and pulse rate operations, as represented in a pulse rate calculation module 410.

In general, the demodulation operation separates the red and infrared signals from the composite signal and removes

-74-

the 625 Hz carrier frequency, leaving raw data points. The raw data points are provided at 625 Hz intervals to the decimation operation which reduces the samples by an order of 10 to samples at 62.5 Hz. The decimation operation also provides some filtering on the samples. The resulting data is subjected to statistics and to the saturation transform operations in order to calculate a saturation value which is very tolerant to motion artifacts and other noise in the signal. The saturation value is ascertained in the saturation calculation module 408, and a pulse rate and a clean plethysmographic waveform is obtained through the pulse rate module 410. Additional detail regarding the various operations is provided in connection with Figures 15-21.

Figure 15 illustrates the operation of the demodulation module 400. The modulated signal format is depicted in Figure 15. One full 625 Hz cycle of the composite signal is depicted in Figure 15 with the first quarter cycle being the active red light plus ambient light signal, the second quarter cycle being an ambient light signal, the third quarter cycle being the active infrared plus ambient light signal, and the fourth quarter cycle being an ambient light signal. As depicted in Figure 15, with a 20 KHz sampling frequency, the single full cycle at 625 Hz described above comprises 32 samples of 20 KHz data, eight samples relating to red plus ambient light, eight samples relating to ambient light, eight samples relating to infrared plus ambient light, and finally eight samples related to ambient light.

Because the signal processing system 334 controls the activation of the light emitters 300, 302, the entire system is synchronous. The data is synchronously divided (and thereby demodulated) into four 8-sample packets, with a time division demultiplexing operation as represented in a demultiplexing module 421. One eight-sample packet 422 represents the red plus ambient light signal; a second eight-sample packet 424 represents an ambient light signal; a third eight-sample packet 426 represents the attenuated infrared

-75-

light plus ambient light signal; and a fourth eight-sample packet 428 represents the ambient light signal. A select signal synchronously controls the demultiplexing operation so as to divide the time-division multiplexed composite signal at the input of the demultiplexer 421 into its four subparts.

A sum of the last four samples from each packet is then calculated, as represented in the summing operations 430, 432, 434, 436 of Figure 15. In the present embodiment, the last four samples are used because a low pass filter in the analog to digital converter 356 of the present embodiment has a settling time. Thus, collecting the last four samples from each 8-sample packet allows the previous signal to clear. This summing operation provides an integration operation which enhances noise immunity. The sum of the respective ambient light samples is then subtracted from the sum of the red and infrared samples, as represented in the subtraction modules 438, 440. The subtraction operation provides some attenuation of the ambient light signal present in the data. In the present embodiment, it has been found that approximately 20 dB attenuation of the ambient light is provided by the operations of the subtraction modules 438, 440. The resultant red and infrared sum values are divided by four, as represented in the divide by four modules 442, 444. Each resultant value provides one sample each of the red and infrared signals at 625 Hz.

It should be understood that the 625 Hz carrier frequency has been removed by the demodulation operation 400. The 625 Hz sample data at the output of the demodulation operation 400 is sample data without the carrier frequency. In order to satisfy Nyquist sampling requirements, less than 20 Hz is needed (understanding that the human pulse is about 25 to 250 beats per minute, or about .4 Hz - 4 Hz). Accordingly, the 625 Hz resolution is reduced to 62.5 Hz in the decimation operation.

Figure 16 illustrates the operations of the decimation module 402. The red and infrared sample data is provided at

-76-

625 Hz to respective red and infrared buffer/filters 450, 452. In the present embodiment, the red and infrared buffer/filters are 519 samples deep. Advantageously, the buffer filters 450, 452 function as continuous first-in, first-out buffers. The 519 samples are subjected to low-pass filtering. Preferably, the low-pass filtering has a cutoff frequency of approximately 7.5 Hz with attenuation of approximately -110 dB. The buffer/filters 450, 452 form a Finite Impulse Response (FIR) filter with coefficients for 519 taps. In order to reduce the sample frequency by ten, the low-pass filter calculation is performed every ten samples, as represented in respective red and infrared decimation by 10 modules 454, 456. In other words, with the transfer of each new ten samples into the buffer/filters 450, 452, a new low pass filter calculation is performed by multiplying the impulse response (coefficients) by the 519 filter taps. Each filter calculation provides one output sample for respective red and infrared output buffers 458, 460. In the present embodiment, the red and infrared output buffers 458, 460 are also continuous FIFO buffers that hold 570 samples of data. The 570 samples provide respective infrared and red samples or packets (also denoted "snapshot" herein) of samples. As depicted in Figure 14, the output buffers provide sample data for the statistics operation module 404, saturation transform module 406, and the pulse rate module 410.

Figure 17 illustrates additional functional operation details of the statistics module 404. In summary, the statistics module 404 provides first order oximetry calculations and RMS signal values for the red and infrared channels. The statistics module also provides a cross-correlation output which indicates a cross-correlation between the red and infrared signals.

As represented in Figure 17, the statistics operation accepts two packets of samples (e.g., 570 samples at 62.5 Hz in the present embodiment) representing the attenuated

-77-

infrared and red signals, with the carrier frequency removed. The respective packets for infrared and red signals are normalized with a log function, as represented in the Log modules 480, 482. The normalization is followed by removal

5   of the DC portion of the signals, as represented in the DC Removal modules 484, 486. In the present embodiment, the DC removal involves ascertaining the DC value of the first one of the samples (or the mean of the first several or the mean of an entire snapshot) from each of the respective red and

10   infrared snapshots, and removing this DC value from all samples in the respective packets.

Once the DC signal is removed, the signals are subjected to bandpass filtering, as represented in red and infrared Bandpass Filter modules 488, 490. In the present embodiment,

15   with 570 samples in each packet, the bandpass filters are configured with 301 taps to provide a FIR filter with a linear phase response and little or no distortion. In the present embodiment, the bandpass filter has a pass band from 34 beats/minute to 250 beats/minute. The 301 taps slide over

20   the 570 samples in order to obtain 270 filtered samples representing the filtered red signal and 270 filtered samples representing the filtered infrared signal. In an ideal case, the bandpass filters 488, 490 remove the DC in the signal. However, the DC removal operations 484, 486 assist in DC

25   removal in the present embodiment.

After filtering, the last 120 samples from each packet (of now 270 samples in the present embodiment) are selected for further processing as represented in Select Last 120 Samples modules 492, 494. The last 120 samples are selected

30   because, in the present embodiment, the first 150 samples fall within the settling time for the Saturation Transfer module 406, which processes the same data packets, as further discussed below.

Conventional saturation equation calculations are

35   performed on the red and infrared 120-sample packets. In the present embodiment, the conventional saturation calculations

are performed in two different ways. For one calculation, the 120-sample packets are processed to obtain their overall RMS value, as represented in the first red and infrared RMS modules 496, 498. The resultant RMS values for red and infrared signals provide input values to a first RED_RMS/IR_RMS ratio operation 500, which provides the RMS red value to RMS infrared value ratio as an input to a saturation equation module 502. As well understood in the art, the ratio of the intensity of red to infrared attenuated light as detected for known red and infrared wavelengths (typically λ_{red} = 650 nm and λ_{IR} = 910 nm) relates to the oxygen saturation of the patient. Accordingly, the saturation equation module 502 represents a conventional look-up table or the like which, for predetermined ratios, provides known saturation values at its output 504. The red and infrared RMS values are also provided as outputs of the statistics operations module 404.

In addition to the conventional saturation operation 502, the 120-sample packets are subjected to a cross-correlation operation as represented in a first cross-correlation module 506. The first cross-correlation module 506 determines if good correlation exists between the infrared and red signals. This cross correlation is advantageous for detecting defective or otherwise malfunctioning detectors. The cross correlation is also advantageous in detecting when the signal model (i.e., the model of Equations (1) - (3)) is satisfied. If correlation becomes too low between the two channels, the signal model is not met. In order to determine this, the normalized cross correlation can be computed by the cross-correlation module 506 for each snapshot of data. One such correlation function is as follows:

$$\frac{\sum S_1 S_2}{\sqrt{\sum S_1^2 \sum S_2^2}}$$

-79-

If the cross correlation is too low, the oximeter 299 provides a warning (e.g., audible, visual, etc.) to the operator. In the present embodiment, if a selected snapshot yields a normalized correlation of less than 0.75, the snapshot does not qualify. Signals which satisfy the signal model will have a correlation greater than the threshold.

The red and infrared 120-sample packets are also subjected to a second saturation operation and cross correlation in the same manner as described above, except the 120 samples are divided into 5 equal bins of samples (i.e., 5 bins of 24 samples each). The RMS, ratio, saturation, and cross correlation operations are performed on a bin-by-bin basis. These operations are represented in the Divide Into Five Equal Bins modules 510, 512, the second red and infrared RMS modules 514, 516, the second RED-RMS/IR-RMS ratio module 518, the second saturation equation module 520 and the second cross correlation module 522 in Figure 17.

Figure 18 illustrates additional detail regarding the saturation transform module 406 depicted in Figure 14. As illustrated in Figure 18, the saturation transform module 406 comprises a reference processor 530, a correlation canceler 531, a master power curve module 554, and a bin power curve module 533. The saturation transform module 406 can be correlated to Figure 7a which has a reference processor 26 and a correlation canceler 27 and an integrator 29 to provide a power curve for separate signal coefficients as depicted in Figure 7c. The saturation transform module 406 obtains a saturation spectrum from the snapshots of data. In other words, the saturation transform 406 provides information of the saturation values present in the snapshots.

As depicted in Figure 18, the reference processor 530 for the saturation transform module 406 has a saturation equation module 532, a reference generator module 534, a DC removal module 536 and a bandpass filter module 538. The red and infrared 570-sample packets from the decimation operation are provided to the reference processor 530. In addition, a

plurality of possible saturation values (the "saturation axis scan") are provided as input to the saturation reference processor 530. In the present embodiment, 117 saturation values are provided as the saturation axis scan. In a preferred embodiment, the 117 saturation values range uniformly from a blood oxygen saturation of 34.8 to 105.0. Accordingly, in the present embodiment, the 117 saturation values provide an axis scan for the reference processor 530 which generates a reference signal for use by the correlation canceler 531. In other words, the reference processor is provided with each of the saturation values, and a resultant reference signal is generated corresponding to the saturation value. The correlation canceler is formed by a joint process estimator 550 and a low pass filter 552 in the present embodiment.

It should be understood that the scan values could be chosen to provide higher or lower resolution than 117 scan values. The scan values could also be non-uniformly spaced.

As illustrated in Figure 18, the saturation equation module 532 accepts the saturation axis scan values as an input and provides a ratio "r_n" as an output. In comparison to the general discussion of Figure 7a-7c, this ratio "r_n" corresponds to the plurality of scan value discussed above in general. The saturation equation simply provides a known ratio "r" (red/infrared) corresponding to the saturation value received as an input.

The ratio "r_n" is provided as an input to the reference generator 534, as are the red and infrared sample packets. The reference generator 534 multiplies either the red or infrared samples by the ratio "r_n" and subtracts the value from the infrared or red samples, respectively. For instance, in the present embodiment, the reference generator 534 multiplies the red samples by the ratio "r_n" and subtracts this value from the infrared samples. The resulting values become the output of the reference generator 534. This operation is completed for each of the saturation

-81-

scan values (e.g., 117 possible values in the present embodiment). Accordingly, the resultant data can be described as 117 reference signal vectors of 570 data points each, hereinafter referred to as the reference signal vectors. This data can be stored in an array or the like.

In other words, assuming that the red and infrared sample packets represent the red $S_{red}(t)$ and infrared $S_{IR}(t)$ measured signals which have primary $s(t)$ and secondary $n(t)$ signal portions, the output of the reference generator becomes the secondary reference signal $n'(t)$, which complies with the signal model defined above, as follows:

$$n'(t) = S_{ir}(t) - r_n S_{red}(t)$$

In the present embodiment, the reference signal vectors and the infrared signal are provided as input to the DC removal module 536 of the reference processor 530. The DC removal module 536, like the DC removal modules 484, 486 in the statistics module 404, ascertains the DC value of the first of the samples for the respective inputs (or mean of the first several or all samples in a packet) and subtracts the respective DC baseline from the sample values. The resulting sample values are subjected to a bandpass filter 538.

The bandpass filter 538 of the reference processor 530 performs the same type of filtering as the bandpass filters 488, 490 of the statistics module 404. Accordingly, each set of 570 samples subjected to bandpass filtering results in 270 remaining samples. The resulting data at a first output 542 of the bandpass filter 538 is one vector of 270 samples (representing the filtered infrared signal in the present embodiment). The resulting data at a second output 540 of the bandpass filter 538, therefore, is 117 reference signal vectors of 270 data points each, corresponding to each of the saturation axis scan values provided to the saturation reference processor 530.

It should be understood that the red and infrared sample packets may be switched in their use in the reference processor 530. In addition, it should be understood that the DC removal module 536 and the bandpass filter module 538 can be executed prior to input of the data to the reference processor 530 because the calculations performed in the reference processor are linear. This results in a significant processing economy.

The outputs of the reference processor 530 provide first and second inputs to a joint process estimator 550 of the type described above with reference to Figure 8. The first input to the joint process estimator 550 is the 270-sample packet representing the infrared signal in the present embodiment. This signal contains primary and secondary signal portions. The second input to the joint process estimator is the 117 reference signal vectors of 270 samples each.

The joint process estimator also receives a lambda input 543, a minimum error input 544 and a number of cells configuration input 545. These parameters are well understood in the art. The lambda parameter is often called the "forgetting parameter" for a joint process estimator. The lambda input 543 provides control for the rate of cancellation for the joint process estimator. In the present embodiment, lambda is set to a low value such as 0.8. Because statistics of the signal are non-stationary, a low value improves tracking. The minimum error input 544 provides an initialization parameter (conventionally known as the "initialization value") for the joint process estimator 550. In the present embodiment, the minimum error value is 10^{-6}. This initialization parameter prevents the joint process estimator 500 from dividing by zero upon initial calculations. The number of cells input 545 to the joint process estimator 550 configures the number of cells for the joint process estimator. In the present embodiment, the number of cells for the saturation transform operation 406 is

-83-

six. As well understood in the art, for each sine wave, the joint process estimator requires two cells. If there are two sine waves in the 35 - 250 beats/minute range, six cells allows for the two heart beat sine waves and one noise sine wave.

The joint process estimator 550 subjects the first input vector on the first input 542 to a correlation cancellation based upon each of the plurality of reference signal vectors provided in the second input 540 to the correlation canceler 531 (all 117 reference vectors in sequence in the present embodiment). The correlation cancellation results in a single output vector for each of the 117 reference vectors. Each output vector represents the information that the first input vector and the corresponding reference signal vector do not have in common. The resulting output vectors are provided as an output to the joint process estimator, and subjected to the low pass filter module 552. In the present embodiment, the low pass filter 552 comprises a FIR filter with 25 taps and with a corner frequency of 10 Hz with the sampling frequency of 62.5 Hz (i.e., at the decimation frequency).

The joint process estimator 550 of the present embodiment has a settling time of 150 data points. Therefore, the last 120 data points from each 270 point output vector are used for further processing. In the present embodiment, the output vectors are further processed together as a whole, and are divided into a plurality of bins of equal number of data points. As depicted in Figure 18, the output vectors are provided to a master power curve module 554 and to a Divide into five Equal Bins module 556. The Divide into Five Equal Bins module 556 divides each of the output vectors into five bins of equal number of data points (e.g., with 120 data points per vector, each bin has 24 data points). Each bin is then provided to the Bin Power Curves module 558.

The Master Power Curve module 554 performs a saturation transform as follows: for each output vector, the sum of the squares of the data points is ascertained. This provides a sum of squares value corresponding to each output vector (each output vector corresponding to one of the saturation scan values). These values provide the basis for a master power curve 555, as further represented in Figure 22. The horizontal axis of the power curve represents the saturation axis scan values and the vertical axis represents the sum of squares value (or output energy) for each output vector. In other words, as depicted in Figure 22, each of the sum of squares could be plotted with the magnitude of the sum of squares value plotted on the vertical "energy output" axis at the point on the horizontal axis of the corresponding saturation scan value which generated that output vector. This results in a master power curve 558, an example of which is depicted in Figure 22. This provides a saturation transform in which the spectral content of the attenuated energy is examined by looking at every possible saturation value and examining the output value for the assumed saturation value. As will be understood, where the first and second inputs to the correlation canceler 531 are mostly correlated, the sum of squares for the corresponding output vector of the correlation canceler 531 will be very low. Conversely, where the correlation between the first and second inputs to the correlation canceler 531 are not significantly correlated, the sum of squares of the output vector will be high. Accordingly, where the spectral content of the reference signal and the first input to the correlation canceler are made up mostly of physiological (e.g., movement of venous blood due to respiration) and non-physiological (e.g., motion induced) noise, the output energy will be low. Where the spectral content of the reference signal and the first input to the correlation canceler are not correlated, the output energy will be much higher.

A corresponding transform is completed by the Bin Power Curves module 558, except a saturation transform power curve is generated for each bin. The resulting power curves are provided as the outputs of the saturation transform module 406.

In general, in accordance with the signal model of the present invention, there will be two peaks in the power curves, as depicted in Figure 22. One peak corresponds to the arterial oxygen saturation of the blood, and one peak corresponds to the venous oxygen concentration of the blood. With reference to the signal model of the present invention, the peak corresponding to the highest saturation value (not necessarily the peak with the greatest magnitude) corresponds to the proportionality coefficient r_a. In other words, the proportionality coefficient r_a corresponds to the red/infrared ratio which will be measured for the arterial saturation. Similarly, peak that corresponds to the lowest saturation value (not necessarily the peak with the lowest magnitude) will generally correspond to the venous oxygen saturation, which corresponds to the proportionality coefficient r_v in the signal model of the present invention. Therefore, the proportionality coefficient r_v will be a red/infrared ratio corresponding to the venous oxygen saturation.

In order to obtain arterial oxygen saturation, the peak in the power curves corresponding to the highest saturation value could be selected. However, to improve confidence in the value, further processing is completed. Figure 19 illustrates the operation of the saturation calculation module 408 based upon the output of the saturation transform module 406 and the output of the statistics module 404. As depicted in figure 19, the bin power curves and the bin statistics are provided to the saturation calculation module 408. In the present embodiment, the master power curves are not provided to the saturation module 408 but can be displayed for a visual check on system operation. The bin

-86-

statistics contain the red and infrared RMS values, the seed
saturation value, and a value representing the cross-
correlation between the red and infrared signals from the
statistics module 404.

5   The saturation calculation module 408 first determines
a plurality of bin attributes as represented by the Compute
Bin Attributes module 560. The Compute Bin Attributes module
560 collects a data bin from the information from the bin
power curves and the information from the bin statistics. In
10   the present embodiment, this operation involves placing the
saturation value of the peak from each power curve
corresponding to the highest saturation value in the data
bin. In the present embodiment, the selection of the highest
peak is performed by first computing the first derivative of
15   the power curve in question by convolving the power curve
with a smoothing differentiator filter function. In the
present embodiment, the smoothing differentiator filter
function (using a FIR filter) has the following coefficients:
     0.014964670230367
20   0.098294046682706   . . .
     0.204468276324813
     2.717182664241813
     5.704485606695227
     0.000000000000000
25   -5.704482606695227
   -2.717182664241813
   -0.204468276324813
   -0.098294046682706   . . .
   -0.014964670230367
30   This filter performs the differentiation and smoothing.
Next, each point in the original power curve in question is
evaluated and determined to be a possible peak if the
following conditions are met: (1) the point is at least 2% of
the maximum value in the power curve; (2) the value of the
35   first derivative changes from greater than zero to less than
or equal to zero. For each point that is found to be a

possible peak, the neighboring points are examined and the largest of the three points is considered to be the true peak.

The peak width for these selected peaks is also calculated. The peak width of a power curve in question is computed by summing all the points in the power curve and subtracting the product of the minimum value in the power curve and the number of points in the power curve. In the present embodiment, the peak width calculation is applied to each of the bin power curves. The maximum value is selected as the peak width.

In addition, the infrared RMS value from the entire snapshot, the red RMS value, the seed saturation value for each bin, and the cross correlation between the red and infrared signals from the statistics module 404 are also placed in the data bin. The attributes are then used to determine whether the data bin consists of acceptable data, as represented in a Bin Qualifying Logic module 562.

If the correlation between the red and infrared signals is too low, the bin is discarded. If the saturation value of the selected peak for a given bin is lower than the seed saturation for the same bin, the peak is replaced with the seed saturation value. If either red or infrared RMS value is below a very small threshold, the bins are all discarded, and no saturation value is provided, because the measured signals are considered to be too small to obtain meaningful data. If no bins contain acceptable data, the exception handling module 563 provides a message to the display 336 that the data is erroneous.

If some bins qualify, those bins that qualify as having acceptable data are selected, and those that do not qualify are replaced with the average of the bins that are accepted. Each bin is given a time stamp in order to maintain the time sequence. A voter operation 565 examines each of the bins and selects the three highest saturation values. These values are forwarded to a clip and smooth operation 566.

The clip and smooth operation 566 basically performs averaging with a low pass filter. The low pass filter provides adjustable smoothing as selected by a Select Smoothing Filter module 568. The Select Smoothing Filter

5 module 568 performs its operation based upon a confidence determination performed by a High Confidence Test module 570. The high confidence test is an examination of the peak width for the bin power curves. The width of the peaks provides some indication of motion by the patient -- wider peaks

10 indicating motion. Therefore, if the peaks are wide, the smoothing filter is slowed down. If peaks are narrow, the smoothing filter speed is increased. Accordingly, the smoothing filter 566 is adjusted based on the confidence level. The output of the clip and smooth module 566 provides

15 the oxygen saturation values in accordance with the present invention.

In the presently preferred embodiment, the clip and smooth filter 566 takes each new saturation value and compares it to the current saturation value. If the

20 magnitude of the difference is less than 16 (percent oxygen saturation) then the value is pass. Otherwise, if the new saturation value is less than the filtered saturation value, the new saturation value is changed to 16 less than the filtered saturation value. If the new saturation value is

25 greater than the filtered saturation value, then the new saturation value is changed to 16 more than the filtered saturation value.

During high confidence (no motion), the smoothing filter is a simple one-pole or exponential smoothing filter which is

30 computed as follows:

$$y(n) = 0.6 * x(n) + 0.4 * y(n-1)$$

where $x(n)$ is the clipped new saturation value, and $y(n)$ is the filtered saturation value.

During motion condition, a three-pole IIR (infinite

35 impulse response) filter is used. Its characteristics are controlled by three time constants t_a, t_b, and t_c with values

-89-

of 0.985, 0.900, and 0.94 respectively. The coefficients for a direct form I, IIR filter are computed from these time constants using the following relationships:

$$a_0 = 0$$

5
$$a_1 = t_b + (t_c)(t_a + t_b)$$

$$a_2 = (-t_b)(t_c)(t_a + t_b + (t_c)(t_a))$$

$$a_3 = (t_b)^2(t_c)^2(t_a)$$

$$b_0 = 1 - t_b - (t_c)(t_a + (t_c)(t_b))$$

10
$$b_1 = 2(t_b)(t_c)(t_a - 1)$$

$$b_2 = (t_b)(t_c)(t_b + (t_c)(t_a) - (t_b)(t_c)(t_a) - t_a)$$

Figure 20 and 21 illustrate the pulse rate module 410
(Figure 14) in greater detail. As illustrated in Figure 20,

15   the heart rate module 410 has a transient removal and
bandpass filter module 578, a motion artifact suppression
module 580, a saturation equation module 582, a motion status
module 584, first and second spectral estimation modules 586,
588, a spectrum analysis module 590, a slew rate limiting

20   module 592, an output filter 594, and an output filter
coefficient module 596.

As further depicted in Figure 20, the heart rate module
410 accepts the infrared and red 570-sample snapshots from
the output of the decimation module 402. The heart rate

25   module 410 further accepts the saturation value which is
output from the saturation calculation module 408.   In
addition, the maximum peak width as calculated by the
confidence test module 570 (same as peak width calculation
described above) is also provided as an input to the heart

30   rate module 410. The infrared and red sample packets, the
saturation value and the output of the motion status module
584 are provided to the motion artifact suppression module
580.

The average peak width value provides an input to a

35   motion status module 584. In the present embodiment, if the
peaks are wide, this is taken as an indication of motion. If

motion is not detected, spectral estimation on the signals is carried out directly without motion artifact suppression.

In the case of motion, motion artifacts are suppressed using the motion artifact suppression module 580. The motion artifact suppression module 580 is nearly identical to the saturation transform module 406. The motion artifact suppression module 580 provides an output which connects as an input to the second spectral estimation module 588. The first and second spectral estimation modules 586, 588 have outputs which provide inputs to the spectrum analysis module 590. The spectrum analysis module 590 also receives an input which is the output of the motion status module 584. The output of the spectrum analysis module 590 is the initial heart rate determination of the heart rate module 410 and is provided as input to the slew rate limiting module 592. The slew rate limiting module 592 connects to the output filter 594. The output filter 594 also receives an input from the output filter coefficient module 596. The output filter 594 provides the filtered heart rate for the display 336 (Figure 11).

In the case of no motion, one of the signals (the infrared signal in the present embodiment) is subjected to DC removal and bandpass filtering as represented in the DC removal and bandpass filter module 578. The DC removal and bandpass filter module 578 provide the same filtering as the DC removal and bandpass filter modules 536, 538. During no motion conditions, the filtered infrared signal is provided to the first spectral estimation module 586.

In the present embodiment, the spectral estimation comprises a Chirp Z transform that provides a frequency spectrum of heart rate information. The Chirp Z transform is used rather than a conventional Fourier Transform because a frequency range for the desired output can be designated in a Chirp Z transform. Accordingly, in the present embodiment, a frequency spectrum of the heart rate is provided between 30 and 250 beats/minute. In the present embodiment, the

frequency spectrum is provided to a spectrum analysis module 590 which selects the first harmonic from the spectrum as the pulse rate. Usually, the first harmonic is the peak in the frequency spectrum that has the greatest magnitude and represents the pulse rate. However, in certain conditions, the second or third harmonic can exhibit the greater magnitude. With this understanding, in order to select the first harmonic, the first peak that has an amplitude of at least 1/20th of the largest peak in the spectrum is selected). This minimizes the possibility of selecting as the heart rate a peak in the Chirp Z transform caused by noise.

In the case of motion, a motion artifact suppression is completed on the snapshot with the motion artifact suppression module 580. The motion artifact suppression module 580 is depicted in greater detail in Figure 21. As can be seen in Figure 21, the motion artifact suppression module 580 is nearly identical to the saturation transform module 406 (Figure 18). Accordingly, the motion artifact suppression module has a motion artifact reference processor 570 and a motion artifact correlation canceler 571.

The motion artifact reference processor 570 is the same as the reference processor 530 of the saturation transform module 406. However, the reference processor 570 utilizes the saturation value from the saturation module 408, rather than completing an entire saturation transform with the 117 saturation scan values. The reference processor 570, therefore, has a saturation equation module 581, a reference generator 582, a DC removal module 583, and a bandpass filter module 585. These modules are the same as corresponding modules in the saturation transform reference processor 530. In the present embodiment, the saturation equation module 581 receives the arterial saturation value from the saturation calculation module 408 rather than doing a saturation axis scan as in the saturation transform module 406. This is because the arterial saturation has been selected, and there

-92-

is no need to perform an axis scan. Accordingly, the output of the saturation equation module 581 corresponds to the proportionality constant r_a (i.e., the expected red to infrared ratio for the arterial saturation value). Otherwise, the reference processor 570 performs the same function as the reference processor 530 of the saturation transform module 406.

The motion artifact correlation canceler 571 is also similar to the saturation transform correlation canceler 531 (Figure 18). However, the motion artifact suppression correlation canceler 571 uses a slightly different motion artifact joint process estimator 572. Accordingly, the motion artifact suppression correlation canceler 571 has a joint process estimator 572 and a low-pass filter 573. The motion artifact joint process estimator 572 differs from the saturation transform joint process estimator 550 in that there are a different number of cells (between 6 and 10 in the present embodiment), as selected by the Number of Cells input 574, in that the forgetting parameter differs (.98 in the present embodiment), and in that the time delay due to adaptation differs. The low-pass filter 573 is the same as the low pass filter 552 of the saturation transform correlation canceler 531.

Because only one saturation value is provided to the reference processor, only one output vector of 270 samples results at the output of the motion artifact suppression correlation canceler 571 for each input packet of 570 samples. In the present embodiment, where the infrared wavelength is provided as a first input to the correlation canceler, the output of the correlation canceler 571 provides a clean infrared waveform. It should be understood that, as described above, the infrared and red wavelength signals could be switched such that a clean red waveform is provided at the output of the motion artifact suppression correlation canceler 571. The output of the correlation canceler 571 is a clean waveform because the actual saturation value of the

patient is known which allows the reference processor 570 to generate a noise reference for inputting to the correlation canceler 571 as the reference signal. The clean waveform at the output of the motion artifact suppression module 580 is a clean plethysmograph waveform which can be forwarded to the display 336.

As described above, an alternative joint process estimator uses the QRD least squares lattice approach (Figure 8a, 9a and 10a). Accordingly, the joint process estimator 573 (as well as the joint process estimator 550) could be replaced with a joint process estimator executing the QRD least squares lattice operation.

Figure 21a depicts an alternative embodiment of the motion artifact suppression module with a joint process estimator 572a replacing the joint process estimator 572. The joint process estimator 572a comprises a QRD least squares lattice system as in Figure 10a. In accordance with this embodiment, different initialization parameters are used as necessary for the QRD algorithm.

The initialization parameters are referenced in Figure 21a as "Number of Cells," "Lambda," "MinSumErr," "GamsInit," and "SumErrInit." Number of Cells and Lambda correspond to like parameters in the joint process estimator 572. GamsInit corresponds to the γ initialization variable for all stages except the zero order stage, which as set forth in the QRD equations above is initialized to '1'. SummErrInit provides the δ initialization parameter referenced above in the QRD equations. In order to avoid overflow, the larger of the actual calculated denominator in each division in the QRD equations and MinSumErr is used. In the present embodiment, the preferred initialization parameters are as follows:

Number of Cells = 6

Lambda = .8

MinSumErr = 10^{-20}

GamsInit = 10^{-2}

SumErrInit = 10^{-6}.

The clean waveform output from the motion artifact suppression module 580 also provides an input to the second spectral estimation module 588. The second spectral estimation module 588 performs the same Chirp Z transform as the first spectral estimation module 586. In the case of no motion, the output from the first spectral estimation module 586 is provided to the spectrum analysis module 586; in the case of motion, the output from the second spectral estimation module 588 is provided to a spectrum analysis module 590. The spectrum analysis module 590 examines the frequency spectrum from the appropriate spectral estimation module to determine the pulse rate. In the case of motion, the spectrum analysis module 590 selects the peak in the spectrum with the highest amplitude, because the motion artifact suppression module 580 attenuates all other frequencies to a value below the actual heart rate peak. In the case of no motion, the spectrum analysis module selects the first harmonic in the spectrum as the heart rate as described above.

The output of the spectrum analysis module 590 provides the raw heart rate as an input to the slew rate limiting module 592, which provides an input to an output filter 594. In the present embodiment, the slew rate limiting module 592 prevents changes greater that 20 beats/minute per 2 second interval.

The output filter 594 comprises an exponential smoothing filter similar to the exponential smoothing filter described above with respect to the clip and smooth filter 566. The output filter is controlled via an output filter coefficient module 596. If motion is large, this filter is slowed down, if there is little or no motion, this filter can sample much faster and still maintain a clean value. The output from the output filter 594 is the pulse of the patient, which is advantageously provided to the display 336.

Alternative To Saturation Transform Module -- Bank Of Filters

An alternative to the saturation transform of the saturation transform module 406 can be implemented with a bank of filters as depicted in Figure 23. As seen in Figure 23, two banks of filters, a first filter bank 600 and a second filter bank 602 are provided. The first filter bank 600 receives a first measured signal $S_{\lambda b}(t)$ (the infrared signal samples in the present embodiment) on a corresponding first filter bank input 604, and the second filter bank 602 receives a second measured signal $S_{\lambda a}(t)$ (the red samples in the present embodiment) on a corresponding second filter bank input 606. In a preferred embodiment, the first and second filter banks utilize static recursive polyphase bandpass filters with fixed center frequencies and corner frequencies. Recursive polyphase filters are described in an article Harris, et. al. "Digital Signal Processing With Efficient Polyphase Recursive All-Pass filters" attached hereto as Appendix A. However, adaptive implementations are also possible. In the present implementation, the recursive polyphase bandpass filter elements are each designed to include a specific center frequency and bandwidth.

There are N filter elements in each filter bank. Each of the filter elements in the first filter bank 600 have a matching (i.e., same center frequency and bandwidth) filter element in the second filter bank 602. The center frequencies and the corner frequencies of N elements are each designed to occupy N frequency ranges, 0 to F_1, F_1 - F_2, F_2 - F_3, F_3 - F_4 . . . F_{N-1} - F_N as shown in Figure 23..

It should be understood that the number of filter elements can range from 1 to infinity. However, in the present embodiment, there are approximately 120 separate filter elements with center frequencies spread evenly across a frequency range of 25 beats/minute - 250 beats/minute.

The outputs of the filters contain information about the primary and secondary signals for the first and second measured signals (red and infrared in the present example) at

the specified frequencies. The outputs for each pair of matching filters (one in the first filter bank 600 and one in the second filter bank 602) are provided to saturation determination modules 610. Figure 23 depicts only one saturation determination module 610 for ease of illustration. However, a saturation determination module can be provided for each matched pair of filter elements for parallel processing. Each saturation determination module has a ratio module 616 and a saturation equation module 618.

The ratio module 616 forms a ratio of the second output to the first output. For instance, in the present example, a ratio of each red RMS value to each corresponding infrared RMS value (Red/IR) is completed in the ratio module 616. The output of the ratio module 616 provides an input to the saturation equation module 618 which references a corresponding saturation value for the input ratio.

The output of the saturation equation modules 618 are collected (as represented in the histogram module 620) for each of the matched filter pairs. However, the data collected is initially a function of frequency and saturation. In order to form a saturation transform curve similar to the curve depicted in Figure 22, a histogram or the like is generated as in Figure 24. The horizontal axis represents the saturation value, and the vertical axis represents a summation of the number of points (outputs from the saturation equation modules 618) collected at each saturation value. In other words, if the output of the saturation equation module 618 for ten different matched filter pairs indicates a saturation value of 98%, then a point in the histogram of Figure 24 would reflect a value of 10 at 98% saturation. This results in a curve similar to the saturation transform curve of Figure 22. This operation is completed in the histogram module 620.

The results of the histogram provide a power curve similar to the power curve of Figure 22. Accordingly, the arterial saturation can be calculated from the histogram by

-97-

selecting the peak (greatest number of occurrences in the area of interest) corresponding to the highest saturation value (e.g., the peak 'c' in Figure peaks corresponding to the highest saturation value peak. Similarly, the venous or background saturation can be determined from the histogram by selecting the peak corresponding to the lowest saturation value (e.g., the peak 'd' in Figure 24), in a manner similar to the processing in the saturation calculation module 408.

It should be understood that as an alternative to the histogram, the output saturation (not necessarily a peak in the histogram) corresponding to the highest saturation value could be selected as the arterial saturation with the corresponding ratio representing r_a. Similarly, the output saturation corresponding to the lowest saturation value could be selected as the venous or background saturation with the corresponding ratio representing r_v. For example, in this embodiment, the entry 'a' in the histogram of Figure 24 would be chosen as the arterial saturation and the entry in the histogram 'b' with the lowest saturation value would be chosen as the venous or background saturation.

ALTERNATIVE DETERMINATION OF COEFFICIENTS r_a AND r_v

As explained above, in accordance with the present invention, primary and secondary signal portions, particularly for pulse oximetry, can be modeled as follows:

$$S_{red} = s_1 + n_1 \qquad (\text{red}) \qquad (89)$$

$$S_{IR} = s_2 + n_2 \qquad (\text{infrared}) \qquad (90)$$

$$s_1 = r_a s_2 \text{ and } n_1 = r_v n_2 \qquad (91)$$

Substituting Equation (91) into Equation (89) provides the following:

$$S_{red} = r_a s_2 + r_v n_2 \qquad (\text{red}) \qquad (92)$$

Note that S_{red} and S_{IR} are used in the model of equations (89) - (92). This is because the discussion below is particularly directed to blood oximetry. S_{red} and S_{IR} correspond to S_1 and S_2 in the preceding text, and the

discussion that follows could be generalized for any measure signal S_1 and S_2.

As explained above, determining r_a and r_v (which correspond to arterial and venous blood oxygen saturation via a saturation equation) can be accomplished using the saturation transform described above doing a scan of many possible coefficients. Another method to obtain r_a and r_v based on red and infrared data is to look for r_a and r_v which minimize the correlation between s_k and n_k, assuming s_k is at least somewhat (and preferably substantially) uncorrelated with n_k (where k = 1 or 2). These values can be found by minimizing the following statistical calculation function for k = 2:

$$Correlation(s_2, n_2) = \left| \sum_i s_2(S_{red_i}, S_{IR_i}, r_a, r_v) n_2(S_{red_i}, S_{IR_i}, r_a, r_v) \right| \tag{93}$$

where i represents time.

It should be understood that other correlation functions such as a normalized correlation could also be used.

Minimizing this quantity often provides a unique pair of r_a and r_v if the noise component is uncorrelated to the desired signal component. Minimizing this quantity can be accomplished by solving Equations (90) and (92) for s_2 and n_2, and finding the minimum of the correlation for possible values of r_a and r_v. Solving for s_2 and n_2 provides the following:

$$\begin{bmatrix} S_{red} \\ S_{IR} \end{bmatrix} = \begin{bmatrix} r_a & r_v \\ 1 & 1 \end{bmatrix} \begin{bmatrix} s_2 \\ n_2 \end{bmatrix}$$

inverting the two-by-two matrix provides:

Thus,

$$\begin{bmatrix} r_a & r_v \\ 1 & 1 \end{bmatrix}^{-1} = \frac{1}{r_a - r_v} \begin{bmatrix} 1 & -r_v \\ -1 & r_a \end{bmatrix}$$

$$\begin{bmatrix} S_2 \\ n_2 \end{bmatrix} = \frac{1}{r_a - r_v} \begin{bmatrix} 1 & -r_v \\ -1 & r_a \end{bmatrix} \begin{bmatrix} S_{red} \\ S_{IR} \end{bmatrix}$$

or:

$$S_2 = \frac{1}{r_a - r_v} (S_{red} - r_v S_{IR})$$

$$n_2 = \frac{1}{r_a - r_v} (-S_{red} + r_a S_{IR})$$

Preferably, the correlation of equation (93) is enhanced with a user specified window function as follows:

$$Correlation(s_2, n_2) = \left| \sum_{i=1}^{N} w_i S_2(S_{red_i}, S_{IR_i}, r_a, r_v) n_2(S_{red_i}, S_{IR_i}, r_a, r_v) \right|$$

$$(93a)$$

The Blackman Window is the presently preferred embodiment. It should be understood that there are many additional functions which minimize the correlation between signal and noise. The function above is simply one. Thus,

$$Correlation(s_2, n_2) = \left| \sum_{i=1}^{N} \left[\frac{w_i}{(r_a - r_v)^2} (S_{red_i} - r_v S_{IR_i}) (-S_{red_i} - r_a S_{IR_i}) \right] \right|$$

$$= \frac{1}{(r_a - r_v)^2} \left| -\sum_{i=1}^{N} (S_{red_i})^2 w_i + (r_a + r_v) \sum_{i=1}^{N} S_{IR_i} S_{red_i} w_i + \sum_{i=1}^{N} (S_{IR_i})^2 w_i \right|$$

$$(93b)$$

In order to implement the minimization on a plurality of discrete data points, the sum of the squares of the red sample points, the sum of the squares of the infrared sample points, and the sum of the product of the red times the infrared sample points are first calculated (including the window function, w_i):

$$RR = \sum_{i=1}^{N} (S_{red_i})^2 w_i$$

$$II = \sum_{i=1}^{N} (S_{IR_i})^2 w_i$$

$$IRR = \sum_{i=1}^{N} (S_{IR}) (S_{red_i}) w_i$$

These values are used in the correlation equation (93b). Thus, the correlation equation becomes an equation in terms of two variables, r_a and r_v. To obtain r_a and r_v, an exhaustive scan is executed for a good cross-section of possible values for r_a and r_v (e.g., 20-50 values each corresponding to saturation values ranging from 30 - 105). The minimum of the correlation function is then selected and the values of r_a and r_v which resulted in the minimum are chosen as r_a and r_v.

Once r_a and r_v have been obtained, arterial oxygen saturation and venous oxygen saturation can be determined by provided r_a and r_v to a saturation equation, such as the saturation equation 502 of the statistics module 404 which provides an oxygen saturation value corresponding to the ratios r_a and r_v.

In a further implementation to obtain r_a and r_v, the same signal model set forth above is again used. In order to determine r_a and r_v in accordance with this implementation, the energy in the signal s_2 is maximized under the constraint that s_2 is uncorrelated with n_2. Again, this implementation

is based upon minimizing the correlation between s and n and on the signal model of the present invention where the signal s relates to the arterial pulse and the signal n is the noise (containing information on the venous blood, as well as motion artifacts and other noise); r_a is the ratio (RED/IR) related to arterial saturation and r_v is the ratio (RED/IR) related to venous saturation. Accordingly, in this implementation of the present invention, r_a and r_v are determined such that the energy of the signal s_2 is maximized where s_2 and n_2 are uncorrelated. The energy of the signal s_2 is given by the following equation:

$$ENERGY(s_2) = \frac{1}{(r_a - r_v)^2} \sum_{i=1}^{N} (S_{red} - r_v S_{IR})^2 \qquad (94)$$

$$= \frac{1}{(r_a - r_v)^2} \sum_{i=1}^{N} (S_{red_i}^2) - 2r_v \sum_{i=1}^{N} (S_{red_i} S_{IR_i}) + r_v^2 \sum_{i=1}^{N} (S_{IR_i}^2) \qquad (95)$$

$$= \frac{1}{(r_a - r_v)^2} [R_1 - 2r_v R_{1,2} + r_v^2 R_2] \qquad (96)$$

where R_1 is the energy of the red signal, R_2 is the energy of the infrared signal and $R_{1,2}$ is the correlation between the red and infrared signals.

The correlation between s_2 and n_2 is given by

$$Correlation(s_2, n_2) = \frac{1}{(r_a - r_v)^2} \sum_{i=1}^{N} (S_{red_i} - r_v S_{ir_i})(-S_{red_i} + r_a S_{IR_i})$$

$$= \frac{1}{(r_a - r_v)^2} [-R_1 + (r_a + r_v) R_{1,2} - r_v r_a R_2] \qquad (97)$$

As explained above, the constraint is that s_k and n_k (k = 2 for the present example) are uncorrelated. This

-102-

"decorrelation constraint" is obtained by setting the correlation of Equation (97) to zero as follows:

$$-R_1 + (r_a + r_v) \, R_{12} - r_a r_v R_2 = 0 \qquad (98)$$

In other words, the goal is to maximize equation (94) under the constraint of equation (98).

In order to obtain the goal, a cost function is defined (e.g., a Lagrangian optimization in the present embodiment) as follows:

$$(99)$$

$$J(r_a, r_v, \mu) = \frac{1}{(r_a - r_v)^2} \left[R_1 - 2 r_v R_{1.2} + r_v^2 R_2 + \mu \left(-R_1 + (r_a + r_v) R_{1.2} - r_a r_v R_2 \right) \right]$$

where μ is the Lagrange multiplier. Finding the value of r_a, r_v and μ that solve the cost function can be accomplished using a constrained optimization method such as described in Luenberger, <u>Linear & Nonlinear Programming</u>, Addison-Wesley, 2d Ed., 1984.

Along the same lines, if we assume that the red and infrared signals S_{red} and S_{IR} are non-static, the functions R_1, R_2 and R_{12} defined above are time dependent. Accordingly, with two equations, two unknowns can be obtained by expressing the decorrelation constraint set forth in equation (98) at two different times. The decorrelation constraint can be expressed at two different times, t_1 and t_2, as follows:

$$-R_1(t_1) + (r_a + r_v) R_{12}(t_1) - r_a r_v R_2(t_1) = 0 \qquad (100)$$

$$-R_1(t_2) + (r_a + r_v) R_{12}(t_2) - r_a r_v R_2(t_2) = 0 \qquad (101)$$

Because equations (100) and (101) are non-linear in r_a and r_v, a change of variables allows the use of linear techniques to solve these two equations. Accordingly, with $x = r_a + r_v$; $y = r_a r_v$ equations (100) and (101) become

$$R_{12}(t_1)x - R_2(t_1)y = R_1(t_1) \qquad (102)$$

$$R_{12}(t_2)x - R_2(t_2)y = R_1(t_2) \qquad (103)$$

5

These equation (102) and (103) can be solved for x and y. Then, solving for r_a and r_v from the changes of variables equations provides the following:

$$r_v + \frac{y}{r_v} = x \quad => \quad r_v^2 - xr_v + y = 0 \qquad (104)$$

10 Solving equation (104) results in two values for r_v. In the present embodiment, the r_v value that results in $x^2 - r_v y > 0$ is selected. If both values of r_v result in $x^2 - r_v y > 0$, the r_v that maximizes the energy of s_2 (Energy(s_2)) at t_2 is selected. r_v is then substituted into the equations above to obtain r_a. Alternatively r_a can be found directly in the same

15 manner r_v was determined.

Alternative To Saturation Transform - Complex FFT

The blood oxygen saturation, pulse rate and a clean plethysmographic waveform of a patient can also be obtained

20 using the signal model of the present invention using a complex FFT, as explained further with reference to Figures 25A-25C. In general, by utilizing the signal model of equations (89)-(92) with two measured signals, each with a first portion and a second portion, where the first portion

25 represents a desired portion of the signal and the second portion represents the undesired portion of the signal, and where the measured signals can be correlated with coefficients r_a and r_v, a fast saturation transform on a

discrete basis can be used on the sample points from the output of the decimation operation 402.

Figure 25A corresponds generally to Figure 14, with the fast saturation transform replacing the previously described saturation transform. In other words, the operations of Figure 25A can replace the operations of Figure 14. As depicted in Figure 25A, the fast saturation transform is represented in a fast saturation transform/pulse rate calculation module 630. As in Figure 14, the outputs are arterial oxygen saturation, a clean plethysmographic waveform, and pulse rate. Figure 25B and 25C illustrate additional detail regarding the fast saturation transform/pulse rate calculation module 630. As depicted in Figure 25B, the fast saturation transform module 620 has infrared log and red log modules 640, 642 to perform a log normalization as in the infrared and red log modules 480, 482 of Figure 17. Similarly, there are infrared DC removal and red DC removal modules 644, 646. In addition, there are infrared and red high-pass filter modules 645, 647, window function modules 648, 640, complex FFT modules 652, 654, select modules 653, 655, magnitude modules 656, 658, threshold modules 660, 662, a point-by-point ratio module 670, a saturation equation module 672, and a select saturation module 680. There are also phase modules 690, 692, a phase difference module 694, and a phase threshold module 696. The output of the select saturation module 680 provides the arterial saturation on an arterial saturation output line 682.

In this alternative embodiment, the snapshot for red and infrared signals is 562 samples from the decimation module 402. The infrared DC removal module 644 and the red DC removal module 646 are slightly different from the infrared and red DC removal modules 484, 486 of Figure 17. In the infrared and red DC removal modules 644, 646 of Figure 25B, the mean of all 563 sample points for each respective channel is calculated. This mean is then removed from each

individual sample point in the respective snapshot in order to remove the baseline DC from each sample. The outputs of the infrared and red DC removal modules 644, 646 provide inputs to respective infrared high-pass filter module 645 and red high-pass filter module 647.

The high-pass filter modules 645, 647 comprise FIR filters with 51 taps for coefficients. Preferably, the high-pass filters comprise Chebychev filters with a side-lobe level parameter of 30 and a corner frequency of .5 Hz (i.e., 30 beats/minute). It will be understood that this filter could be varied for performance. With 562 sample points entering the high-pass filters, and with 51 taps for coefficients, there are 512 samples provided from these respective infrared and red snapshots at the output of the high-pass filter modules. The output of the high-pass filter modules provides an input to the window function modules 648, 650 for each respective channel.

The window function modules 648, 650 perform a conventional windowing function. A Kaiser windowing function is used in the present embodiment. The functions throughout Figure 25B maintain a point-by-point analysis. In the present embodiment, the time bandwidth product for the Kaiser window function is 7. The output of the window function modules provides an input to the respective complex Fast Fourier Transform (FFT) modules 652, 654.

The complex FFT modules 652, 654 perform complex FFTs on respective infrared and red channels on the data snapshots. The data from the complex FFTs is then analyzed in two paths, once which examines the magnitude and one which examines the phase from the complex FFT data points. However, prior to further processing, the data is provided to respective infrared and red select modules 653, 655 because the output of the FFT operation will provide repetitive information from 0-1/2 the sampling rate and from 1/2 the sampling rate to the sampling rate. The select modules select only samples from 0-1/2 the sampling rate (e.g., 0-31.25 Hz in the present

embodiment) and then select from those samples to cover a frequency range of the heart rate and one or more harmonics of the heart rate. In the present embodiment, samples which fall in the frequency range of 20 beats per minute to 500

5   beats per minute are selected. This value can be varied in order to obtain harmonics of the heart rate as desired. Accordingly, the output of the select modules results in less than 256 samples. In the present embodiment, the sample points 2-68 of the outputs of the FFTs are utilized for

10   further processing.

In the first path of processing, the output from the select modules 653, 655 are provided to respective infrared and red magnitude modules 656, 658. The magnitude modules 656, 658 perform a magnitude function wherein the magnitude

15   on a point-by-point basis of the complex FFT points is selected for each of the respective channels. The outputs of the magnitude modules 656, 658 provide an input to infrared and red threshold modules 660, 662.

The threshold modules 660, 662 examine the sample

20   points, on a point-by-point basis, to select those points where the magnitude of an individual point is above a particular threshold which is set at a percentage of the maximum magnitude detected among all the remaining points in the snapshots. In the present embodiment, the percentage for

25   the threshold operation is selected as 1% of the maximum magnitude.

After thresholding, the data points are forwarded to a point-by-point ratio module 670. The point-by-point ratio module takes the red over infrared ratio of the values on a

30   point-by-point basis. However, a further test is performed to qualify the points for which a ratio is taken. As seen in Figure 25B, the sample points output from the select modules 653, 655 are also provided to infrared and red phase modules 690, 692. The phase modules 690, 692 select the phase value

35   from the complex FFT points. The output of the phase modules

690, 692 is then presented to a phase difference module 694.

The phase difference module 694 calculates the difference in phase between the corresponding data points from the phase modules 690, 692. If the magnitude of the phase difference between any two corresponding points is less than a particular threshold (e.g., 0.1 radians) in the present embodiment), then the sample points qualify. If the phase of two corresponding sample points is too far apart, then the sample points are not used. The output of the phase threshold module 696 provides an enable input to the RED/IR rate module 670. Accordingly, in order for the ratio of a particular pair of sample points to be taken, the three tests are executed:

6.   the red sample must pass the red threshold 660;

7.   the infrared sample must pass the infrared threshold 662; and

8.   the phase between the two points must be less than the predefined threshold as determined in the phase threshold 696.

For those sample points which qualify, a ratio is taken in the ratio module 670. For those points which do not qualify, the saturation is set to zero at the output of the saturation equation 672.

The resulting ratios are provided to a saturation equation module which is the same as the saturation equation modules 502, 520 in the statistics module 504. In other words, the saturation equation module 672 accepts the ratio on a point-by-point basis and provides as an output a corresponding saturation value corresponding to the discrete ratio points. The saturation points output from the saturation equation module 672 provide a series of saturation points which could be plotted as saturation with respect to frequency. The frequency reference was entered into the points at the complex FFT stage.

The arterial (and the venous) saturation can then be selected, as represented in the select arterial saturation module 680, in one of two methods according to the present invention. According to one method, the arterial saturation value can be selected simply as the point corresponding to the largest saturation value for all points output from the saturation equation module 672 for a packet. Alternatively, a histogram similar to the histogram of Figure 22 can be generated in which the number of saturation values at different frequencies (points) are summed to form a histogram of the number of occurrences for each particular saturation value. In either method, the arterial saturation can be obtained and provided as an output to the select arterial saturation module on the arterial saturation output line 682. In order to obtain the venous saturation, the minimum arterial saturation value, of points that exhibit non-zero value, is selected rather than the maximum arterial saturation value. The saturation can be provided to the display 336.

The fast saturation transform information can also be used to provide the pulse rate and the clean plethysmographic wave form as further illustrated in Figure 25C. In order to obtain the pulse rate and a clean plethysmographic wave form, several additional functions are necessary. As seen in Figure 25C, the pulse rate and clean plethysmographic wave form are determined using a window function module 700, a spectrum analysis module 702 and an inverse window function module 704.

As depicted in Figure 25C, the input to the window function module 700 is obtained from the output of the complex FFT modules 652 or 654. In the present embodiment, only one measured signal is necessary. Another input to the window function module 700 is the arterial saturation obtained from the output of the select arterial saturation module 680.

The window function module performs a windowing function selected to pass those frequencies that significantly correlate to the frequencies which exhibited saturation values very close to the arterial saturation value. In the present embodiment, the following windowing function is selected:

$$1 - \left[\frac{SAT_{art} - SAT_n}{100}\right]^{15} \qquad (105)$$

where SAT_n equals the saturation value corresponding to each particular frequency for the sample points and SAT_{art} represents the arterial saturation as chosen at the output of the select arterial saturation module 680. This window function is applied to the window function input representing the complex FFT of either the red or the infrared signal. The output of the window function module 700 is a red or infrared signal represented with a frequency spectrum as determined by the FFT, with motion artifacts removed by the windowing function. It should be understood that many possible window functions can be provided. In addition, with the window function described above, it should be understood that using a higher power will provide more noise suppression.

In order to obtain pulse rate, the output points from the window function module 700 are provided to a spectrum analysis module 702. The spectrum analysis module 702 is the same as the spectrum analysis module 590 of Figure 20. In other words, the spectrum analysis module 702 determines the pulse rate by determining the first harmonic in the frequency spectrum represented by the output points of the windowing function 700. The output of spectrum analysis module 702 is the pulse rate.

In order to obtain a clean plethysmographic waveform, the output of the windowing function 700 is applied to an inverse window function module 704. The inverse window

function module 704 completes an inverse of the Kaiser window function of the window function module 648 or 650 of Figure 25B. In other words, the inverse window function 704 does a point-by-point inverse of the Kaiser function for points that are still defined. The output is a clean plethysmographic waveform.

Accordingly, by using a complex FFT and windowing functions, the noise can be suppressed from the plethysmographic waveform in order to obtain the arterial saturation, the pulse rate, and a clean plethysmographic waveform. It should be understood that although the above description relates to operations primarily in the frequency domain, operations that obtain similar results could also be accomplished in the time domain.

Relation to Generalized Equations

The measurements described for pulse oximetry above are now related back to the more generalized discussion above. The signals (logarithm converted) transmitted through the finger 310 at each wavelength λa and λb are:

$$S_{\lambda a}(t) = S_{\lambda red1}(t) = \epsilon_{HbO2,\lambda a} c^{A}_{HbO2} x^{A}(t) + \epsilon_{Hb,\lambda a} c^{A}_{Hb} x^{A}(t)$$

$$+ \epsilon_{HbO2,\lambda a} c^{V}_{HbO2} x^{V}(t) + \epsilon_{Hb,\lambda a} c^{V}_{Hb} x^{V}(t) + n_{\lambda a}(t); \qquad (105a)$$

$$S_{\lambda a}(t) = \epsilon_{HbO2,\lambda a} c^{A}_{HbO2} x^{A}(t) + \epsilon_{Hb,\lambda a} c^{A}_{Hb} x^{A}(t) + n_{\lambda a}(t); \qquad (105b)$$

$$S_{\lambda a}(t) = s_{\lambda a}(t) + n_{\lambda a}(t); \qquad (105c)$$

$$S_{\lambda b}(t) = S_{\lambda red2}(t) = \epsilon_{HbO2,\lambda b} c^{A}_{HbO2} x^{A}(t) + \epsilon_{Hb,\lambda b} c^{A}_{Hb} x^{A}(t)$$

$$+ \epsilon_{HbO2,\lambda b} c^{V}_{HbO2} x^{V}(t) + \epsilon_{Hb,\lambda b} c^{V}_{Hb} x^{V}(t) + n_{\lambda b}(t); \qquad (106a)$$

$$S_{\lambda b}(t) = \epsilon_{HbO2,\lambda b} c^{A}_{HbO2} x^{A}(t) + \epsilon_{Hb,\lambda b} c^{A}_{Hb} x^{A}(t) + n_{\lambda b}(t) \qquad (106b)$$

$$S_{\lambda b}(t) = s_{\lambda b}(t) + n_{\lambda b}(t) \qquad (106c)$$

The variables above are best understood as correlated to Figure 6c as follows: assume the layer in figure 6c containing A_3 and A_4 represents venous blood in the test medium, with A_3 representing deoxygenated hemoglobin (Hb) and

5   A_4 representing oxygenated hemoglobin (HBO2) in the venous blood. Similarly, assume that the layer in Figure 6c containing A_5 and A_6 represents arterial blood in the test medium, with A_5 representing deoxygenated hemoglobin (Hb) and A_6 representing oxygenated hemoglobin (HBO2) in the arterial

10   blood. Accordingly, $c^V HbO2$ represents the concentration of oxygenated hemoglobin in the venous blood, $c^V Hb$ represents the concentration of deoxygenated hemoglobin in the venous blood, x^V represents the thickness of the venous blood (e.g., the thickness the layer containing A_3 and A_4). Similarly,

15     $c^A HbO2$ represents the concentration of oxygenated hemoglobin in the arterial blood, $c^A Hb$ represents the concentration of deoxygenated hemoglobin in the arterial blood, and x^A represents the thickness of the arterial blood (e.g., the thickness of the layer containing A_5 and A_6)

20   The wavelengths chosen are typically one in the visible red range, i.e., λa, and one in the infrared range, i.e., λb. Typical wavelength values chosen are $\lambda a = 660\text{nm}$ and $\lambda b = 910\text{nm}$. In accordance with the constant saturation method, it is assumed that $c^A_{HbO2}(t)/c^A_{Hb}(t) = \text{constant}_1$ and $c^V_{HbO2}(t)/c^V_{Hb}(t)$

25     $= \text{constant}_2$. The oxygen saturation of arterial and venous blood changes slowly, if at all, with respect to the sample rate, making this a valid assumption. The proportionality coefficients for equations (105) and (106) can then be written as:

30

$$r_a(t) = \frac{\epsilon_{HbO2,\lambda a} c^A_{HbO2} x(t) + \epsilon_{Hb,\lambda a} c_{Hb} x(t)}{\epsilon_{HbO2,\lambda b} c^A_{HbO2} x(t) + \epsilon_{Hb,\lambda b} c^A_{Hb} x(t)} \qquad (107)$$

$$s_{\lambda a}(t) = r_a(t) s_{\lambda b}(t) \qquad (108a)$$

-112-

$$n_{\lambda a}(t) \neq r_a(t) n_{\lambda b}(t) \qquad (109a)$$

$$n_{\lambda a}(t) = r_v(t) n_{\lambda b}(t) \qquad (108b)$$

$$s_{\lambda a}(t) \neq r_v(t) s_{\lambda b}(t) \qquad (109b)$$

In pulse oximetry, it is typically the case that both equations (108) and (109) can be satisfied simultaneously.

Multiplying equation (106) by $r_a(t)$ and then subtracting equation (106) from equation (105), a non-zero secondary reference signal n'(t) is determined by:

$$n'(t) = S_{\lambda a}(t) - r_a(t) S_{\lambda b}(t) \qquad (110a)$$

$$= \epsilon_{HbO2,\lambda a} c^V_{HbO2} x^V(t) + \epsilon_{Hb,\lambda a} c^V_{Hb} x^V(t) + n_{\lambda a}(t)$$

$$- r_a(t) [\epsilon_{HbO2,\lambda b} c^V_{HbO2} x^V(t) + \epsilon_{Hb,\lambda b} c^V_{Hb} x^V(t) + n_{\lambda b}(t)]. \qquad (111a)$$

Multiplying equation (106) by $r_v(t)$ and then subtracting equation (106) from equation (105), a non-zero primary reference signal s'(t) is determined by:

$$s'(t) = S_{\lambda a}(t) - r_v(t) S_{\lambda b}(t) \qquad (110b)$$

$$= s_{\lambda a}(t) - r_v(t) s_{\lambda b}(t) \qquad (111b)$$

The constant saturation assumption does not cause the venous contribution to the absorption to be canceled along with the primary signal portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$. Thus, frequencies associated with both the low frequency modulated absorption due to venous absorption when the patient is still and the modulated absorption due to venous absorption when the patient is moving are represented in the secondary reference signal n'(t). Thus, the correlation canceler or other methods described above remove or derive both erratically modulated absorption due to venous blood in the

finger under motion and the constant low frequency cyclic absorption of venous blood.

To illustrate the operation of the oximeter of Figure 11 to obtain clean waveform, Figures 26 and 27 depict signals measured for input to a reference processor of the present invention which employs the constant saturation method, i.e., the signals $S_{\lambda a}(t) = S_{\lambda red}(t)$ and $S_{\lambda b}(t) = S_{\lambda IR}(t)$. A first segment 26a and 27a of each of the signals is relatively undisturbed by motion artifact, i.e., the patient did not move substantially during the time period in which these segments were measured. These segments 26a and 27a are thus generally representative of the primary plethysmographic waveform at each of the measured wavelengths. A second segment 26b and 27b of each of the signals is affected by motion artifact, i.e., the patient did move during the time period in which these segments were measured. Each of these segments 26b and 27b shows large motion induced excursions in the measured signal. A third segment 26c and 27c of each of the signals is again relatively unaffected by motion artifact and is thus generally representative of the primary plethysmographic waveform at each of the measured wavelengths.

Figure 28 shows the secondary reference signal $n'(t) = n_{\lambda a}(t) - r_a n_{\lambda b}(t)$, as determined by a reference processor of the present invention. Again, the secondary reference signal $n'(t)$ is correlated to the secondary signal portions $n_{\lambda a}$ and $n_{\lambda b}$. Thus, a first segment 28a of the secondary reference signal $n'(t)$ is generally flat, corresponding to the fact that there is very little motion induced noise in the first segments 26a and 27a of each signal. A second segment 28b of the secondary reference signal $n'(t)$ exhibits large excursions, corresponding to the large motion induced excursions in each of the measured signals. A third segment 28c of the noise reference signal $n'(t)$ is generally flat, again corresponding to the lack of motion artifact in the third segments 26c and 27c of each measured signal.

It should also be understood that a reference processor could be utilized in order to obtain the primary reference signal $s'(t) = s_{\lambda a} - r_v s_{\lambda b}(t)$. The primary reference signal $s'(t)$ would be generally indicative of the plethysmograph waveform.

Figures 29 and 30 show the approximations $s''_{\lambda a}(t)$ and $s''_{\lambda b}(t)$ to the primary signals $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ as estimated by a correlation canceler using a secondary reference signal $n'(t)$. Note that the scale of Figures 26 through 30 is not the same for each figure to better illustrate changes in each signal. Figures 29 and 30 illustrate the effect of correlation cancellation using the secondary reference signal $n'(t)$ as determined by the reference processor. Segments 29b and 30b are not dominated by motion induced noise as were segments 26b and 27b of the measured signals. Additionally, segments 29a, 30a, 29c, and 30c have not been substantially changed from the measured signal segments 26a, 27a, 26c, and 27c where there was no motion induced noise.

It should be understood that approximation $n''_{\lambda a}(t)$ and $n''_{\lambda b}(t)$ to the secondary signals $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ as estimated by a correlation canceler using a primary reference signal $s'(t)$ can also be determined in accordance with the present invention.

METHOD FOR ESTIMATING PRIMARY AND SECONDARY
SIGNAL PORTIONS OF
MEASURED SIGNALS IN A PULSE OXIMETER

Implementing the various embodiments of the correlation canceler described above in software is relatively straightforward given the equations set forth above, and the detailed description above. However, a copy of a computer program subroutine, written in the C programming language, which calculates a primary reference $s'(t)$ using the constant saturation method and, using a joint process estimator 572 which implements a joint process estimator using the equations (54) - (64) is set forth in Appendix B. This joint process estimator estimates a good approximation to the

-115-

primary signal portions of two measured signals, each having a primary portion which is correlated to the primary reference signal s'(t) and a secondary portion which is correlated to the secondary reference signal n'(t). This subroutine is another way to implement the steps illustrated in the flowchart of Figure 9 for a monitor particularly adapted for pulse oximetry. The two signals are measured at two different wavelengths λa and λb, where λa is typically in the visible region and λb is typically in the infrared region. For example, in one embodiment of the present invention, tailored specifically to perform pulse oximetry using the constant saturation method, $\lambda a = 660nm$ and $\lambda b = 940nm$.

The correspondence of the program variables to the variables defined in equations (54) - (64) in the discussion of the joint process estimator is as follows:

$$\Delta_m(t) = nc[m].Delta$$
$$\Gamma_{f,m}(t) = nc[m].fref$$
$$\Gamma_{b,m}(t) = nc[m].bref$$
$$f_m(t) = nc[m].ferr$$
$$b_m(t) = nc[m].berr$$
$$\vartheta_m(t) = nc[m].Fswsqr$$
$$\beta_m(t) = nc[m].Bswsqr$$
$$\gamma_m(t) = nc[m].Gamma$$
$$\rho_{m,\lambda a}(t) = nc[m].Roh\_a$$
$$\rho_{m,\lambda b}(t) = nc[m].Roh\_b$$
$$e_{m,\lambda a}(t) = nc[m].err\_a$$
$$e_{m,\lambda b}(t) = nc[m].err\_b$$
$$\kappa_{m,\lambda a}(t) = nc[m].K\_a$$
$$\kappa_{m,\lambda b}(t) = nc[m].K\_b$$

A first portion of the program performs the initialization of the registers 90, 92, 96, and 98 and intermediate variable values as in the "INITIALIZED CORRELATION CANCELER" action block 120. A second portion of

the program performs the time updates of the delay element variables 110 with the value at the input of each delay element variable 110 is stored in the delay element variable 110 as in the "TIME UPDATE OF LEFT [Z^{-1}] ELEMENTS" action

5   block 130. The calculation of saturation is performed in a separate module. Various methods for calculation of the oxygen saturation are known to those skilled in the art. One such calculation is described in the articles by G.A. Mook, et al, and Michael R. Neuman cited above. Once the

10   concentration of oxygenated hemoglobin and deoxygenated hemoglobin are determined, the value of the saturation is determined similarly to equations (72) through (79) wherein measurements at times t_1 and t_2 are made at different, yet proximate times over which the saturation is relatively

15   constant. For pulse oximetry, the average saturation at time $t = (t_1+t_2)/2$ is then determined by:

$$Sat_{arterial}(t) = \frac{c_{HbO2}^A(t)}{c_{HbO2}^A(t) + c_{Hb}^A(t)} \tag{112a}$$

$$= \frac{\epsilon_{Hb,\lambda a} - \epsilon_{Hb,\lambda b}(\Delta s_{\lambda a}/\Delta s_{\lambda b})}{\epsilon_{HB,\lambda a} - \epsilon_{HbO2,\lambda a} - (\epsilon_{HB,\lambda b} - \epsilon_{HbO2,\lambda b})(\Delta s_{\lambda a}/\Delta s_{\lambda b})} \tag{112b}$$

$$Sat_{venous}(t) = \frac{c_{HbO2}^V(t)}{c_{HbO2}^V(t) + c_{Hb}^V(t)} \tag{113a}$$

$$= \frac{\epsilon_{Hb,\lambda a} - \epsilon_{Hb,\lambda b}(\Delta n_{\lambda a}/\Delta n_{\lambda b})}{\epsilon_{HB,\lambda a} - \epsilon_{HbO2,\lambda a} - (\epsilon_{HB,\lambda b} - \epsilon_{HbO2,\lambda b})(\Delta n_{\lambda a}/\Delta n_{\lambda b})} \tag{113b}$$

A third portion of the subroutine calculates the primary

20   reference or secondary reference, as in the "CALCULATE PRIMARY OR SECONDARY REFERENCE ($s'(t)$ or $n'(t)$) FOR TWO MEASURED SIGNAL SAMPLES" action block 140 for the signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ using the proportionality constants $r_a(t)$ and

$r_v(t)$ determined by the constant saturation method as in equation (3). The saturation is calculated in a separate subroutine and a value of $r_a(t)$ or $r_v(t)$ is imported to the present subroutine for estimating either the primary portions $s_{\lambda a}(t)$ and $s_{\lambda b}(t)$ or the secondary portions $n_{\lambda a}(t)$ and $n_{\lambda b}(t)$ of the composite measured signals $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$.

A fourth portion of the program performs Z-stage update as in the "ZERO STAGE UPDATE" action block 150 where the Z-stage forward prediction error $F_0(t)$ and Z-stage backward prediction error $b_0(t)$ are set equal to the value of the reference signal $n'(t)$ or $s'(t)$ just calculated. Additionally zero-stage values of intermediate variables ϑ_0 and $\beta_0(t)$ (nc[m].Fswsqr and nc[m].Bswsqr in the program) are calculated for use in setting registers 90, 92, 96, and 98 values in the least-squares lattice predictor 70 in the regression filters 80a and 80b.

A fifth portion of the program is an iterative loop wherein the loop counter, M, is reset to zero with a maximum of m=NC_CELLS, as in the "m=0" action block 160 in Figure 9. NC_ CELLS is a predetermined maximum value of iterations for the loop. A typical value for NC_CELLS is between 6 and 10, for example. The conditions of the loop are set such that the loop iterates a minimum of five times and continues to iterate until a test for conversion is met or m-NC_ CELLS. The test for conversion is whether or not the sum of the weighted sum of four prediction errors plus the weighted sum of backward prediction errors is less than a small number, typically 0.00001 (i.e., $\vartheta_m(t)+\beta m(t) \leq 0.00001$).

A sixth portion of the program calculates the forward and backward reflection coefficient $\Gamma_{m,f}(t)$ and $\Gamma_{m,b}(t)$ register 90 and 92 values (nc[m].fref and nc[m].bref in the program) as in the "ORDER UPDATE m^{th}-STAGE OF LSL-PREDICTOR" action block 170. Then forward and backward prediction errors $f_m(t)$ and $b_m(t)$ (nc[m].ferr and nc[m].berr in the program) are calculated. Additionally, intermediate variables $\vartheta_m(t)$, $\beta m(t)$, and $\gamma(t)$ (nc[m].Fswsqr, nc[m].Bswsqr,

-118-

nc[m]. gamma in the program) are calculated. The first cycle of the loop uses the value for nc[0].Fswsqr and nc[0].Bswsqr calculated in the ZERO STAGE UPDATE portion of the program.

A seventh portion of the program, still within the loop begun in the fifth portion of the program, calculates the regression coefficient register 96 and 98 values $\kappa_{m,\lambda a}(t)$ and $\kappa_{m,\lambda b}(t)$ (nc[m].K_a and nc[m].K_b in the program) in both regression filters, as in the "ORDER UPDATE m^{th} STAGE OF REGRESSION FILTER(S)" action block 180. Intermediate error signals and variables $e_{m,\lambda a}(t)$, $e_{m,\lambda b}(t)$, $\rho_{m,\lambda a}(t)$, and $\rho_{m,\lambda b}(t)$ (nc[m].err_a and nc[m].err_b, nc[m].roh_a, and nc[m].roh_b in the subroutine) are also calculated.

The loop iterates until the test for convergence is passed. The test for convergence of the joint process estimator is performed each time the loop iterates analogously to the "DONE" action block 190. If the sum of the weighted sums of the forward and backward prediction errors $\Im_m(t) + \beta m(t)$ is less than or equal to 0.00001, the loop terminates. Otherwise, sixth and seventh portions of the program repeat.

The output of the present subroutine is a good approximation to the primary signals $s"_{\lambda a}(t)$ and $s"_{\lambda b}(t)$ or the secondary signals $n"_{\lambda a}(t)$ and $n"_{\lambda b}(t)$ for the set of samples $S_{\lambda a}(t)$ and $S_{\lambda b}(t)$ input to the program. After approximations to the primary signal portions or the secondary signals portions of many sets of measured signal samples are estimated by the joint process estimator, a compilation of the outputs provides waves which are good approximations to the plethysmographic wave or motion artifact at each wavelength, λa and λb.

It should be understood that the subroutine of Appendix B is merely one embodiment which implements the equations (54) - (64). Although implementation of the normalized and QRD-LSL equations is also straightforward, a subroutine for the normalized equations is attached as Appendix C, and a

subroutine for the QRD-LSL algorithm is attached as Appendix D.

While one embodiment of a physiological monitor incorporating a processor of the present invention for determining a reference signal for use in a correlation canceler, such as an adaptive noise canceler, to remove or derive primary and secondary components from a physiological measurement has been described in the form of a pulse oximeter, it will be obvious to one skilled in the art that other types of physiological monitors may also employ the above described techniques.

Furthermore, the signal processing techniques described in the present invention may be used to compute the arterial and venous blood oxygen saturations of a physiological system on a continuous or nearly continuous time basis. These calculations may be performed, regardless of whether or not the physiological system undergoes voluntary motion.

Furthermore, it will be understood that transformations of measured signals other than logarithmic conversion and determination of a proportionality factor which allows removal or derivation of the primary or secondary signal portions for determination of a reference signal are possible. Additionally, although the proportionality factor r has been described herein as a ratio of a portion of a first signal to a portion of a second signal, a similar proportionality constant determined as a ratio of a portion of a second signal to a portion of a first signal could equally well be utilized in the processor of the present invention. In the latter case, a secondary reference signal would generally resemble $n'(t) = n_{\lambda b}(t) - rn_{\lambda a}(t)$.

Furthermore, it will be understood that correlation cancellation techniques other than joint process estimation may be used together with the reference signals of the present invention. These may include but are not limited to least mean square algorithms, wavelet transforms, spectral

estimation techniques, neural networks, Weiner and Kalman filters among others.

One skilled in the art will realize that many different types of physiological monitors may employ the teachings of the present invention. Other types of physiological monitors include, but are in not limited to, electro cardiographs, blood pressure monitors, blood constituent monitors (other than oxygen saturation) monitors, capnographs, heart rate monitors, respiration monitors, or depth of anesthesia monitors. Additionally, monitors which measure the pressure and quantity of a substance within the body such as a breathalizer, a drug monitor, a cholesterol monitor, a glucose monitor, a carbon dioxide monitor, a glucose monitor, or a carbon monoxide monitor may also employ the above-described techniques.

Furthermore, one skilled in the art will realize that the above described techniques of primary or secondary signal removal or derivation from a composite signal including both primary and secondary components can also be performed on electrocardiography (ECG) signals which are derived from positions on the body which are close and highly correlated to each other. It should be understood that a tripolar Laplacian electrode sensor such as that depicted in Figure 31 which is a modification of a bipolar Laplacian electrode sensor discussed in the article "Body Surface Laplacian ECG Mapping" by Bin He and Richard J. Cohen contained in the journal IEEE Transactions on Biomedical Engineering, Vol. 39, No. 11, November 1992 could be used as an ECG sensor. It must also be understood that there are a myriad of possible ECG sensor geometry's that may be used to satisfy the requirements of the present invention. The same type of sensor could also be used for EEG and EMG measurements.

Furthermore, one skilled in the art will realize that the above described techniques can also be performed on signals made up of reflected energy, rather than transmitted energy. One skilled in the art will also realize that a

primary or secondary portion of a measured signal of any type
of energy, including but not limited to sound energy, X-ray
energy, gamma ray energy, or light energy can be estimated by
the techniques described above. Thus, one skilled in the art

5   will realize that the techniques of the present invention can
be applied in such monitors as those using ultrasound where
a signal is transmitted through a portion of the body and
reflected back from within the body back through this portion
of the body.   Additionally, monitors such as echo

10   cardiographs may also utilize the techniques of the present
invention since they too rely on transmission and reflection.

While the present invention has been described in terms
of a physiological monitor, one skilled in the art will
realize that the signal processing techniques of the present

15   invention can be applied in many areas, including but not
limited to the processing of a physiological signal. The
present invention may be applied in any situation where a
signal processor comprising a detector receives a first
signal which includes a first primary signal portion and a

20   first secondary signal portion and a second signal which
includes a second primary signal portion and a second
secondary signal portion. Thus, the signal processor of the
present invention is readily applicable to numerous signal
processing areas.